Luglio 1989
Numero 251
Anno XXII
Volume XLIII

La correttezza del software in applicazioni a rischio

L'obiettivo di produrre software corretto nel caso di applicazioni a rischio rilevante è perseguito razionalizzando tutto il processo di sviluppo, in modo da evitare errori e/o rilevare meglio quelli eventualmente commessi

di Giuseppe Di Marco e Alberto Pasquini

L'enorme progresso dei sistemi basati sull'uso dei calcolatori risponde a una esigenza specifica della nostra civiltà tecnologica che richiede una pronta soluzione di problemi sempre più complessi. I calcolatori hanno così affiancato, o sostituito, l'uomo in compiti ove è richiesta la capacità di prendere decisioni tempestive sulla base di grandi quantità di informazioni. Alcuni di questi compiti, quali per esempio il controllo del traffico aereo, il monitoraggio delle condizioni cliniche di pazienti in reparti di terapia intensiva, il controllo dei processi chimici o nucleari, possono essere definiti a rischio rilevante. È di estrema importanza che, nell'ambito di queste applicazioni, il calcolatore prenda decisioni corrette. Infatti un suo errore potrebbe comportare gravi danni in termini economici, di vite umane o di impatto ambientale.

La verifica della correttezza di un sistema basato su calcolatori tenderà ad accertare che esso si comporti esattamente come richiesto dall'utente cioè, in altre parole, che esegua senza errori il compito per cui è stata richiesta la sua realizzazione. Sebbene questa attività implichi una valutazione del sistema nella sua globalità, i mezzi tecnici attualmente disponibili permettono di verificare in modo sufficientemente preciso solo la sua parte *hardware*. Al contrario la verifica del *software* si è dimostrata un problema di difficile soluzione a causa di alcune sue caratteristiche intrinseche.

Semplificando al massimo si può dire che il software è l'insieme di istruzioni (in realtà di programmi, procedure, dati e documentazione) che descrivono in modo formale e sequenziale le funzioni che il calcolatore deve svolgere. Poiché i calcolatori dell'attuale generazione, salvo poche eccezioni, sono in grado di elaborare le informazioni solo in modo sequenziale, le istruzioni debbono guidarli passo per passo nell'esecuzione dei loro compiti. Come esempio si consideri un sistema basato su calcolatore che debba mantenere un impianto nucleare in condizioni sicure. Una delle sue funzioni è quella di garantire l'integrità del sistema di raffreddamento del reattore, evitando che la portata dell'acqua scenda a livelli critici e che la pressione superi i valori permessi. Questa funzione viene ricondotta a una sequenza di azioni e di decisioni che, come è mostrato nell'illustrazione a pagina 18, consiste: nell'acquisizione dei dati di pressione e di portata provenienti dagli strumenti, nel riconoscimento e nella segnalazione delle condizioni del circuito di raffreddamento e, infine, nell'avviamento delle azioni di emergenza necessarie (allarmi, comandi per l'arresto del reattore, avviamento di pompe e apertura di valvole). Nel caso di applicazioni reali, le funzioni affidate ai sistemi computerizzati sono estremamente complesse; per esempio un tipico sistema, che svolga la funzione di supervisione di processo in una centrale nucleare, acquisisce ed elabora dai 3000 ai 7000 tipi di dati di processo differenti.

Per verificare che il software sia corretto si dovrebbe utilizzare il metodo della verifica formale, che consiste nel cercare di provare il software in un modo analogo a quello usato in matematica per la dimostrazione di teoremi. Ma, nei casi reali, questa prova è di una complessità paragonabile a quella del software che si intende verificare ed è quindi altrettanto soggetta a errori.

Alternativamente si potrebbe ricorrere a una prova esaustiva di tutte le possibili condizioni di impianto previste. Facendo riferimento all'esempio precedente, sarà quindi necessario provare ogni possibile combinazione dei valori di pressione e di portata del circuito di raffreddamento, cioè ogni possibile configurazione degli ingressi al calcolatore. Ma anche questo tipo di verifica è evidentemente impossibile, in quanto richiederebbe l'esecuzione di un'infinità di prove.

Più realisticamente, con la tecnica dell'*input domain partitioning*, si possono raggruppare questi valori negli insiemi che rappresentano i quattro possibili stati del circuito di raffreddamento (pressione e portata normali, pressione normale e bassa portata, alta pressione e portata normale, alta pressione e bassa portata). Ciò comporta l'esecuzione di quattro prove, cioè che venga fornita una coppia di valori di pressione e portata (configurazione degli ingressi) per ognuno degli stati individuati e che vengano verificate le quattro azioni di protezione conseguenti. Purtroppo, appena si passa ad applicazioni reali, gli insiemi che rappresentano i possibili stati del processo sotto controllo divengono difficili da individuare e crescono di numero in modo esponenziale a mano a mano che cresce la complessità della funzione che il calcolatore deve svolgere.

L'impiego dei calcolatori ha rappresentato una risposta tecnologica alla necessità di affiancare o sostituire l'uomo in compiti dove è richiesta la capacità di prendere decisioni tempestive tenendo conto simultaneamente di una grande quantità di dati. Alcuni di questi compiti, quali per esempio il controllo del traffico aereo o dei processi chimici e nucleari, sono definiti a rischio rilevante. L'illustrazione mostra la tipica immagine video che viene presentata all'operatore da un calcolatore con funzioni di controllo di processo di una centrale nucleare. In particolare, essa fornisce informazioni che riguardano la temperatura misurata all'interno di una sezione di un quarto del nocciolo nucleare. Gli elementi di combustibile che lo costituiscono sono rappresentati dai quadrati colorati. Per ognuno di essi viene fornito, tramite un codice di colori, il margine di sicurezza rispetto alla temperatura massima ammissibile. Nell'ambito di questa particolare applicazione è di estrema importanza che il calcolatore dia informazioni corrette. Infatti, un errore nella presentazione della temperatura degli elementi di combustibile del nocciolo potrebbe causare una condizione potenzialmente pericolosa, tale da richiedere l'intervento del sistema di protezione per evitare una eventuale fusione degli elementi di combustibile.

Inoltre eseguire la sequenza di istruzioni corrispondente a una specifica configurazione degli ingressi non è sufficiente ad assicurare che questa sequenza sia corretta: potrebbero infatti esserci al suo interno due errori che in quel caso particolare si annullano a vicenda; oppure degli errori che si manifestano solo per una particolare configurazione degli ingressi diversa da quella provata. Ritornando all'esempio precedente, se fosse stata introdotta una istruzione errata e non prevista del tipo «se la pressione è maggiore di 17 megapascal allora salta alla fine del programma», non ci si renderebbe conto dell'errore fino a quando non venisse fornito come ingresso un valore di pressione superiore a 17 megapascal. Sebbene l'esempio possa sembrare artificioso non è affatto raro che sviste e/o dimenticanze da parte del programmatore conducano a situazioni che, anche se non appaiono altrettanto «maliziose», producono risultati analoghi.

Inoltre questi metodi, generalmente detti di verifica a posteriori, non sono in grado di individuare gli errori dovuti alla difficoltà di definire correttamente e univocamente le funzioni che il software deve svolgere. Infatti nelle applicazioni in tempo reale, e cioè in quelle in cui le informazioni (ingressi) provengano dal mondo reale e non dal progettista, non è possibile prevedere a priori tutte le possibili condizioni a cui il software sarà sottoposto (si veda in proposito l'articolo *Il software per la difesa antimissili balistici* di Herbert Lin in «Le Scienze», n. 210, febbraio 1986). Inoltre quanto più la funzione è complessa e/o ipotetica, come nel caso delle funzioni di sicurezza, tanto più tale definizione si fa difficile e quindi foriera di errori.

Verificare a posteriori se un determinato prodotto software è privo di errori e correggere quelli rilevati è l'approccio alla correttezza generalmente usato. Ma nelle applicazioni ove tali errori dovrebbero essere completamente assenti, quali quelle a rischio rilevante, esso diviene complesso e oneroso e non fornisce sufficienti garanzie sui risultati ottenuti.

Per superare i limiti della verifica a posteriori si ricorre perciò a tutte quelle tecniche che l'esperienza ha dimostrato utili per prevenire gli errori che potrebbero essere introdotti durante lo sviluppo del software.

Lo sviluppo del software comporta il passaggio da una descrizione funzionale dei compiti da realizzare a una sequenza di istruzioni eseguibile da un calcolatore. Tale processo è complesso, in quanto richiede una visione mentale globale e dettagliata del problema e di tutte le trasformazioni necessarie per lo sviluppo. Per esempio la funzione di protezione di un reattore nucleare comporta la necessità di realizzare un «meccanismo» che in base ai dati di processo (temperatura, pressione, portata ecc.) produca una serie di comandi e di indicazioni (allarmi, comandi per l'arresto di emergenza, per l'avviamento di pompe e l'apertura di valvole) tali da mantenerlo in condizioni sicure. Una tipica applicazione industriale di questo genere (quale un sistema a microprocessori che svolge le funzioni di protezione e controllo in una centrale nucleare) ha comportato la stesura di circa 500 000 istruzioni e un carico di lavoro stimabile in 200-300 anni-uomo.

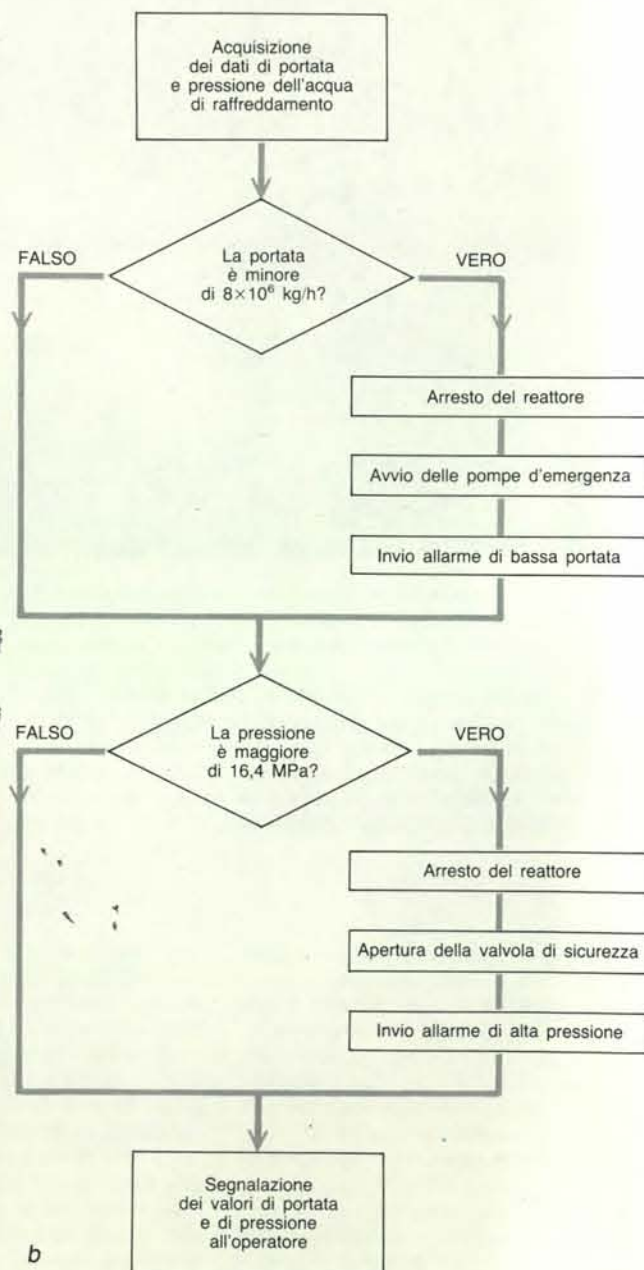
FUNZIONE: GARANTIRE L'INTEGRITÀ DEL SISTEMA DI RAFFREDDAMENTO DI UN REATTORE NUCLEARE

Sulla base dei valori di pressione e portata letti dagli strumenti posizionati sul sistema di raffreddamento:

1) Verificare che ci si trovi in condizioni normali di funzionamento (portata maggiore di 8×10^6 kg/h e pressione minore di 16,4 megapascal); altrimenti intraprendere le azioni di sicurezza opportune (arresto del reattore, allarmi, avvio delle pompe di emergenza, apertura della valvola di sicurezza).
2) Informare l'operatore sullo stato del sistema.

a

```
PROCEDURE: INTEGRITÀ SISTEMA RAFFREDDAMENTO

    INPUT DATI: PORTATA
    INPUT DATI: PRESSIONE

    IF: PORTATA < 8.10E6 kg/h

        THEN

        OUTPUT COMANDO :  ARRESTO REATTORE
        OUTPUT COMANDO :  AVVIO POMPE EMERGENZA
        OUTPUT ALLARME  :  BASSA PORTATA

        ELSE

    END IF

    IF: PRESSIONE > 16.4 MPASCAL

        THEN

        OUTPUT COMANDO :  ARRESTO REATTORE
        OUTPUT COMANDO :  APERTURA VALVOLA SICUREZZA
        OUTPUT ALLARME  :  ALTA PRESSIONE

        ELSE

    END IF

    OUTPUT DATI :  PORTATA
    OUTPUT DATI :  PRESSIONE

END PROCEDURE
```

c

I calcolatori dell'attuale generazione, salvo poche eccezioni, sono in grado di elaborare le informazioni solo in modo sequenziale, perciò debbono essere guidati passo per passo nell'esecuzione dei compiti a essi assegnati. La loro programmazione richiede quindi che la funzione da svolgere sia ricondotta a una sequenza di azioni e decisioni. Questa, a sua volta, viene tradotta in una successione di istruzioni comprensibili ai calcolatori. In *a* viene descritta la funzione predisposta per garantire l'integrità del sistema di raffreddamento di un reattore nucleare. In *b* viene mostrata la sequenza di azioni (*rettangoli*) e decisioni (*rombi*) nella quale la funzione è stata ridotta. In *c* è riportata la traduzione di tale sequenza in uno pseudolinguaggio di programmazione comprensibile al calcolatore.

Si ricorre perciò a una razionalizzazione del processo di sviluppo del software individuando e organizzando gli elementi che contribuiscono alla sua realizzazione. In primo luogo si suddivide il processo in una successione di fasi attraverso le quali, con raffinamenti successivi, si passa dall'idea di un oggetto, che deve svolgere determinati compiti, alla sua realizzazione concreta. Il processo di produzione così definito prende il nome di «ciclo di sviluppo del sistema». Contemporaneamente il personale coinvolto viene organizzato in gruppi di specialisti (gruppo di sviluppo, gruppo di verifica, gruppo della garanzia di qualità e gruppo di gestione) dei quali si individuano compiti e responsabilità. Inoltre, per ciascuna fase dello sviluppo, vanno identificati i documenti e i metodi che i vari gruppi di specialisti devono adottare per controllare e sviluppare il progetto.

Il ciclo di sviluppo del sistema, nella forma attualmente più consolidata, inizia con la fase di individuazione delle funzioni che il sistema deve svolgere: i bisogni dell'utente e/o le caratteristiche del processo da controllare vengono analizzati dal gruppo di sviluppo, cercando di definirli e concordarli, senza ambiguità, sottintesi e tecnicismi, in un documento generalmente chiamato «Requisiti del sistema». Dopo aver identificato le parti hardware e software da realizzare, lo sviluppo procede secondo tre fasi ben precise: durante la prima viene definito *che cosa* ciascuna parte deve fare in un documento di specifiche; durante la seconda viene definito *come* lo deve fare in un documento che descrive il progetto; infine, durante la terza fase, le parti hardware e software vengono realizzate. Riferendosi all'esempio precedente, *a*, *b* e *c* dell'illustrazione nella pagina a fronte mostrano il tipo di risultato che si ottiene alla fine di ognuna di queste fasi.

Attualmente sono state individuate alcune metodologie specifiche che guidano questo sviluppo, con l'obiettivo di prevenire gli errori che possono essere introdotti nelle varie fasi, nonché alcune tecniche che permettono di individuare gli errori eventualmente introdotti nel passaggio da una fase a quella successiva («verifica passo passo»).

La progettazione inizia con una decomposizione della funzione globale che il software deve svolgere, la quale viene ridotta a un insieme di sottofunzioni più semplici e fra loro indipendenti (progettazione *top-down*). Il processo viene ripetuto più volte, fino a ottenere delle sottofunzioni così elementari da essere mentalmente riconducibili a una successione di azioni e decisioni che possono a questo punto essere tradotte in istruzioni per il calcolatore nella fase di codifica. L'illustrazione a pagina 20 schematizza, utilizzando una tecnica ben precisa chiamata SADT (da *Structured Analysis and Design Technique*), questo processo di suddivisione e di raffinamenti successivi.

Oltre agli evidenti vantaggi di razionalizzazione della progettazione, questa tecnica porta alla realizzazione di programmi semplici, modulari e strutturati, e perciò più facili da capire e da verificare. Ognuno di essi è infatti di piccole dimensioni, svolge una sola e specifica sottofunzione, e la sequenza delle sue istruzioni può essere iniziata e terminata solamente in punti ben precisi (programmazione strutturata).

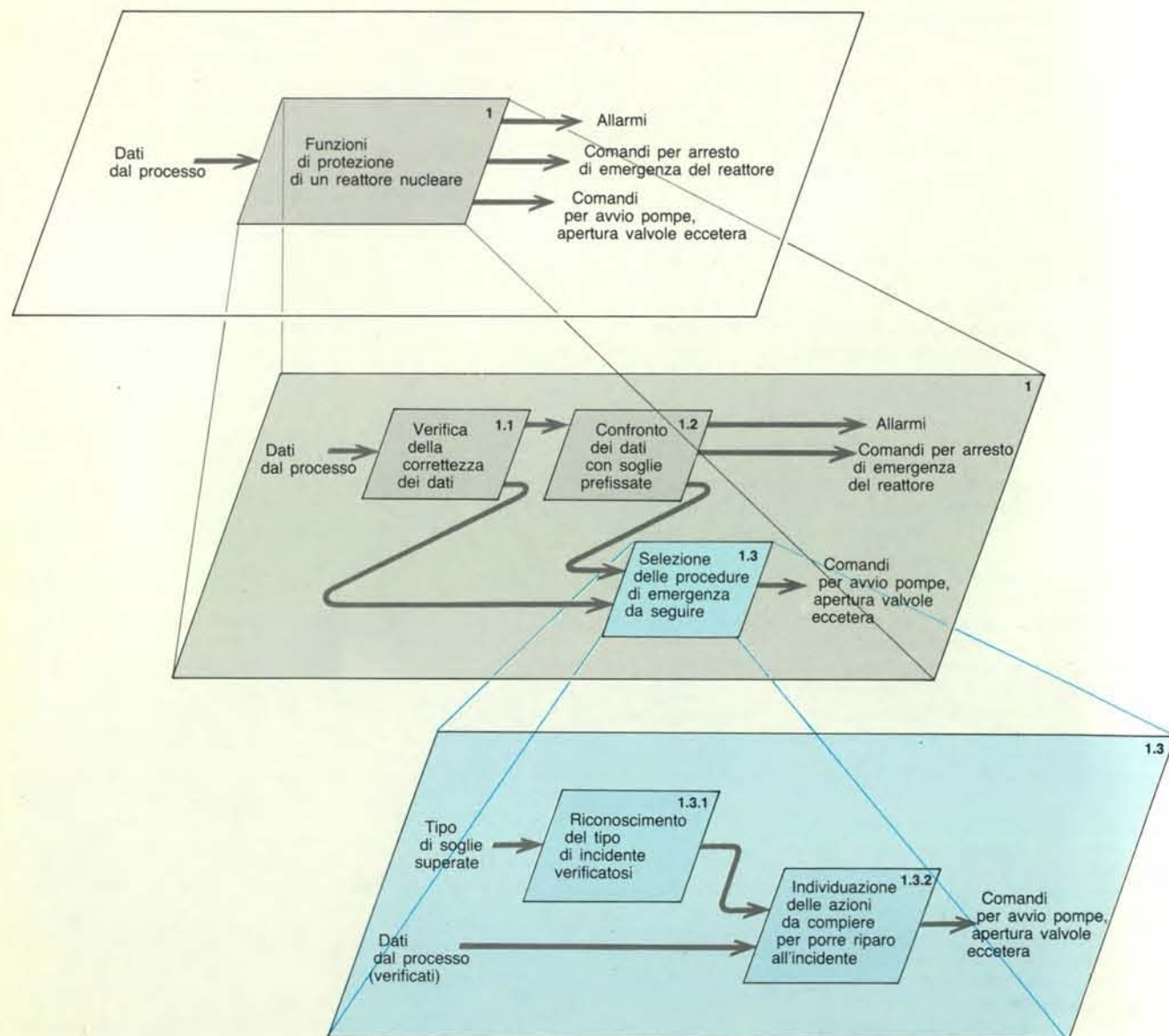
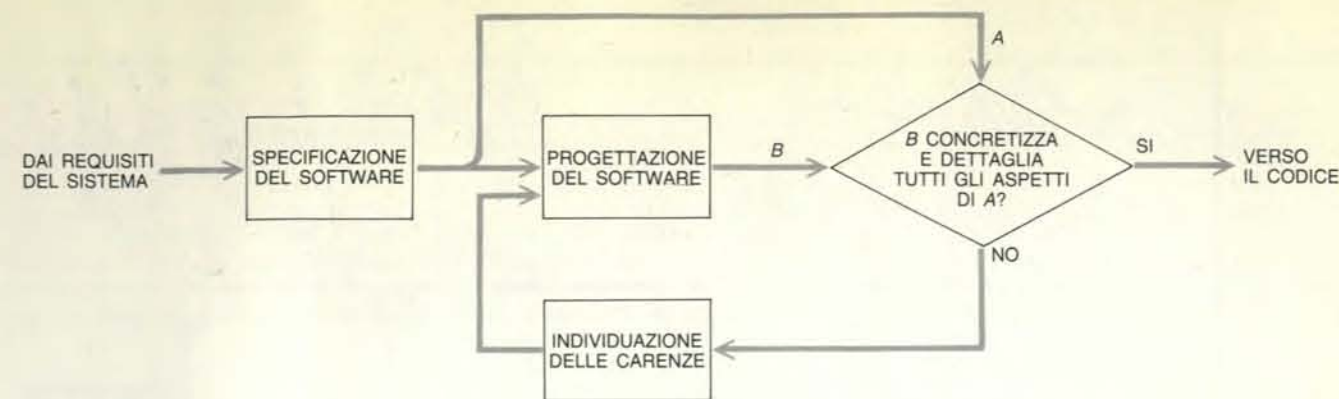
Ovviamente, anche se si adotta una tecnica quale quella sin qui descritta, durante la progettazione del software si possono ugualmente commettere degli errori.

Consideriamo il caso di un microprocessore che guidi un semaforo stradale

Una delle vie adottate nel perseguire l'obiettivo della correttezza del software consiste nella razionalizzazione del processo di sviluppo. Sviluppare un sistema basato su calcolatore significa passare dall'idea di un oggetto destinato a svolgere determinati compiti alla concreta messa a punto di hardware e software in grado di eseguirli. L'illustrazione mostra le fasi fondamentali attraverso cui questo passaggio si realizza (il cosiddetto «ciclo di sviluppo del sistema»). Con il passaggio dai «bisogni dell'utente» alle fasi successive il sistema viene prima definito e poi realizzato e provato nelle sue componenti di hardware e software. Nelle ultime fasi queste due componenti vengono integrate per realizzare il sistema finale, che viene quindi provato nella sua globalità. Il corretto sviluppo del software viene controllato, parallelamente alla sua realizzazione, con una «verifica passo passo».

con l'obiettivo di smaltire il traffico in modo «intelligente». Una delle funzioni individuate nella fase di specifica potrebbe essere quella di dare via libera per prima alla direttrice di traffico sulla quale vi sono più macchine in attesa. In fase di progettazione si potrebbe non tener conto dell'eventualità che il numero delle macchine in attesa sia uguale in entrambe le direzioni e quindi condurre il microprocessore in una situazione di «stallo».

La verifica passo passo tende a individuare questo e altri tipi di errori per permettere di rimuoverli quanto prima possibile senza aspettare di arrivare alle prove finali. Generalmente la verifica viene condotta da un gruppo di persone diverso da quello che ha realizzato il progetto (seguendo le indicazioni del vecchio adagio secondo cui nessuno è buon giudice di stesso). Il gruppo di verifica controlla che la documentazione, prodotta alla fine di ogni fase di sviluppo, sia coerente con la documentazione finale della fase precedente, in quanto ne sono stati dettagliati e concretizzati ulteriormente tutti gli aspetti.

L'obiettivo che ci si pone facendo ricorso alle tecniche di progetto e di verifica sin qui descritte è quello di produrre un software privo di errori, ma questo obiettivo può attualmente essere paragonato a quello del moto perpetuo: non tutto quello che teoricamente è possibile in realtà è ottenibile. Comunque, come nel caso del moto perpetuo, ogni innovazione tecnologica che si avvicini a questo obiettivo comporta innegabili vantaggi pratici, così anche tutti gli sforzi compiuti per la produzione di un software corretto otterranno, quantomeno,

Durante la fase di progettazione la funzione che il software deve svolgere viene tradotta in istruzioni per il calcolatore. Le difficoltà incontrate, nel compiere direttamente questa trasformazione, hanno spinto all'ideazione di metodologie che la guidino attraverso una serie di passi successivi. La metodologia generalmente adottata consiste nel ridurre la funzione in una serie di sottofunzioni più semplici e maggiormente dettagliate. Il processo viene ripetuto più volte fino a ottenere sottofunzioni così elementari da essere mentalmente riconducibili a una serie di azioni e decisioni, che possono a loro volta essere tradotte in istruzioni. L'illustrazione mostra un'applicazione di tale metodologia, mediante la tecnica SADT (da *Structured Analysis and Design Technique*), alla funzione di protezione di un reattore nucleare. Tale funzione, rappresentata nel primo riquadro in alto, viene decomposta in tre sottofunzioni nel riquadro centrale. Una di queste viene ulteriormente decomposta nel riquadro in basso. Le frecce entranti rappresentano le informazioni fornite in ingresso alle funzioni e alle sottofunzioni, mentre le frecce uscenti indicano il risultato delle relative elaborazioni.

La verifica passo passo ha l'obiettivo di individuare precocemente gli errori senza aspettare di arrivare alle prove finali, quando potrebbe essere difficile e costoso rimuoverli completamente. Questa attività viene condotta perciò sin dall'inizio dello sviluppo del software, per evitare che gli errori, commessi durante le prime fasi, si propaghino e amplifichino nel corso del progetto. La documentazione prodotta alla fine di ogni fase viene confrontata con la documentazione finale della fase precedente per verificare che ne siano stati concretizzati e dettagliati ulteriormente tutti gli aspetti. L'illustrazione mostra la sequenza di attività (*rettangoli*) implicate nella verifica della «progettazione del software». Il flusso dei documenti è rappresentato dalle frecce entranti e dalle frecce uscenti; la verifica vera e propria, sulla base della documentazione prodotta dalla «progettazione del software», è rappresentata dal rombo.

il risultato di ridurre il numero degli errori introdotti.

Poiché non è ragionevole ritenere di riuscire a eliminare completamente gli errori, il ciclo di sviluppo prevede una serie di prove finali con cui si cerca di individuare gli errori che sono sfuggiti alla verifica e una serie di prove con cui si controlla che il prodotto, considerato nel suo complesso di hardware e software, soddisfi le richieste del cliente (fase di validazione).

L'illustrazione a pagina 24 mostra, per un'applicazione concreta, il numero degli errori commessi in ogni fase del ciclo di sviluppo del software e in quali fasi si è riusciti a individuarli. L'esempio si riferisce a un processo di sviluppo di software con funzioni di protezione di un reattore nucleare. Tale sviluppo è stato effettuato a scopo sperimentale come parte di un programma di ricerca svolto dal gruppo norvegese Halden Reactor Project.

Nella fase in cui vengono eseguite le prove finali, il sistema, o le sue parti, è sollecitato sistematicamente con l'obiettivo di rilevare gli eventuali errori presenti. Ma una sollecitazione del sistema che sia in grado di rilevarli tutti richiederebbe l'esecuzione di un'infinità di prove; pertanto è necessario ricorrere a una loro selezione, in modo da scegliere quelle che permettono di individuare il maggior numero di errori possibile.

Vi sono due metodi basilari che sono utilizzati nella scelta delle prove da eseguire: un metodo sistemistico («scatola nera») e un metodo strutturale («scatola bianca»).

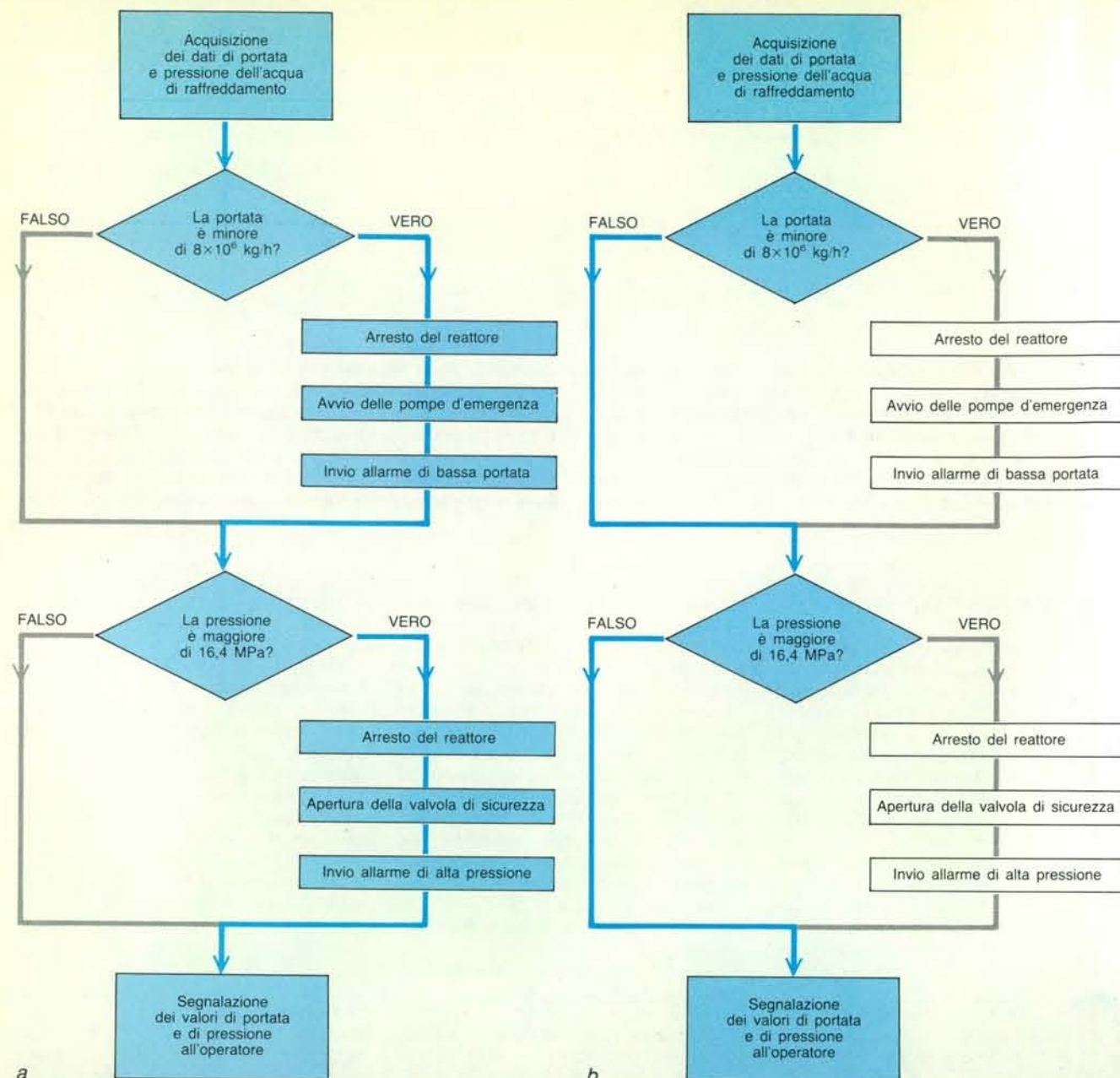
Nel primo caso le prove sono derivate dai requisiti del sistema e/o dalle sue specifiche senza tenere conto della struttura interna del programma: l'obiettivo è quello di attivare tutte le sue funzioni, oppure il maggior numero di ingressi possibili scelti in modo casuale, oppure le condizioni limite del sistema. Un importante ausilio nell'esecuzione di queste prove è dato da strumenti automatici che simulano le condizioni di impianto previste. Nella maggior parte dei casi si tratta di un altro calcolatore che è in grado di riprodurre il comportamento del processo da controllare. Esso analizza i comandi forniti dal programma in prova e gli fornisce, in tempo reale, le stesse risposte (configurazione degli ingressi) che riceverebbe dal processo. Il limite di questi simulatori è costituito dal fatto che, per quanto accurati essi possano essere, non riescono a riprodurre che in modo approssimativo e incompleto il comportamento del processo.

Nel secondo caso le prove sono costruite in modo da sollecitare gli elementi che costituiscono la struttura interna del programma. Tali elementi, in ordine crescente di complessità, sono: le istruzioni, i rami e i cammini.

Sollecitare tutte le istruzioni vuol dire trovare un insieme di dati di ingresso tale che esse siano eseguite tutte almeno una volta, indipendentemente dal loro ordine. Tornando all'esempio del sistema di raffreddamento del reattore, la parte *b* dell'illustrazione di pagina 18 può essere vista come una rappresentazione semplificata della struttura interna del programma. Per eseguire tutte le istruzioni sarà sufficiente una sola prova che abbia come configurazione degli ingressi un valore di portata minore di 8×10^6 kg/h e una pressione maggiore di 16,4 megapascal. La parte di struttura del programma così eseguita è evidenziata nella parte *a* dell'illustrazione di pagina 22. Questo metodo di prova è il più semplice da seguire, ma anche quello che offre minori garanzie sulla correttezza del programma: abbiamo già visto che parecchi tipi di errore possono rimanere non rilevati anche se tutte le istruzioni sono state eseguite.

Maggiori garanzie si ottengono se vengono sollecitati tutti i rami del programma, cioè entrambe le sequenze di istruzioni relative alle possibili scelte di ogni decisione (vero o falso). In accordo con questo metodo è necessario eseguire un'ulteriore prova, in aggiunta a quella che esegue tutte le istruzioni, mostrata nella parte *b* della stessa illustrazione.

Un metodo ancora più efficace dei precedenti consiste nel sollecitare tutti i cammini di un programma, cioè tutte le possibili sequenze di istruzioni che portano dal suo inizio alla sua fine. Tornando all'esempio, nelle parti *c* e *d* dell'illustrazione di pagina 23 si vede come vadano eseguite due ulteriori prove, in aggiunta a quelle già descritte, per ottenere questo scopo. Anche in questo caso si può ricorrere a strumenti automatici che controllino il programma durante l'esecuzione delle prove. Essi sono in grado di determinare quali istruzioni, rami e cammini sono stati eseguiti e in che percentuali, essendo così di aiuto nella individuazione di ulteriori possibili prove.

Confrontando i diversi metodi utilizzati nell'esempio, emerge che il numero di prove da eseguire aumenta con la complessità degli elementi della struttura che si vogliono sollecitare: una prova nel caso delle istruzioni, due nel caso dei rami e quattro nel caso dei cammini. In particolare i cammini possibili raddoppiano a ogni decisione contenuta nel programma; perciò, in applicazioni non banali, il loro numero diviene troppo elevato ed è quindi difficile costruire prove che li eseguano tutti. Inoltre se è possibile affermare che quando il programma è corretto tutte queste prove avranno risultato corretto, non è possibile affermare il contrario. Perciò il me-

Eseguire una prova significa verificare che, data una determinata configurazione in ingresso a un programma, questo fornisca in uscita un risultato esatto. Provare tutte le possibili configurazioni richiederebbe però l'esecuzione di un numero esorbitante di prove. Nella pratica è pertanto necessario ricorrere a una loro selezione, scegliendo quelle prove che offrono maggiori garanzie sulla correttezza del programma, vale a dire quelle che permettono di individuare il maggior numero possibile di eventuali errori. Uno dei criteri attraverso cui si può effettuare la selezione si basa sulla scelta delle prove che più sollecitano gli elementi della struttura interna del programma. Tali elementi, in ordine crescente di complessità, sono: i rami, le istruzioni e i cammini. L'illustrazione mostra le prove selezionate, secondo i criteri suddetti, per il programma descritto nell'illustrazione di pagina 18. Per sollecitare tutte le istruzioni è sufficiente la sola prova *a*, che vengano cioè forniti in ingresso i valori di portata e di pressione elencati nella prima colonna della tabella in basso. Il programma verrà allora eseguito secondo il cammino rappresentato in colore nella parte *a* della figura, cammino che prevede, per l'appunto, l'esecuzione di tutte le istruzioni. Un metodo di prova più completo richiede la sollecitazione di tutti i possibili rami del programma, vale a dire di entrambe le sequenze di istruzioni che seguono i punti di decisione (sia quella sul ramo relativo alla condizione «vero» sia quella sul ramo relativo alla condizione «falso»). Nell'esempio, avendo già provato le condizioni «vero» in *a*, sarà sufficiente un'ulteriore prova per le condizioni «falso». Tale prova è mostrata nella seconda colonna della tabella e il cammino relativo nella parte *b* della figura. Un metodo di prova ancor più efficace dei precedenti consiste nel sollecitare tutti i possibili cammini del programma, cioè nell'eseguire tutte le possibili sequenze di istruzioni che portano dal suo inizio alla sua fine. Allo scopo saranno necessarie due ulteriori prove, in aggiunta alle precedenti, i cui cammini sono mostrati in *c* e *d*.

	PROVA *a*	PROVA *b*	PROVA *c*	PROVA *d*
VALORI DELLA PORTATA IN INGRESSO	MINORE DI 8×10^6 kg/h	MAGGIORE DI 8×10^6 kg/h	MINORE DI 8×10^6 kg/h	MAGGIORE DI 8×10^6 kg/h
VALORI DELLA PRESSIONE IN INGRESSO	MAGGIORE DI 16,4 MPa	MINORE DI 16,4 MPa	MINORE DI 16,4 MPa	MAGGIORE DI 16,4 MPa

todo usualmente eseguito è quello di utilizzare queste prove in modo parziale e in combinazione con quelle selezionate in base al metodo sistemistico.

A partire dagli anni settanta si è tentato di stimare il grado di correttezza raggiunto partendo da dati raccolti durante le prove finali. L'esperienza ha dimostrato che la frequenza di ritrovamento degli errori segue approssimativamente una legge esponenziale decrescente con il tempo. Si è cercato di costruire perciò dei modelli che simulino la storia degli errori e dai quali si possano trarre opportune indicazioni sul numero degli errori residui presenti nel sistema.

Un altro metodo utilizzato si basa sulla considerazione che il software sarà tanto più corretto quanto più le prove sono state efficaci e, mediante modelli matematici che tengano conto di questa efficacia, tenta di stimare il numero di errori residui. Una delle tecniche ricalca quella utilizzata per stimare il numero di pesci in un lago o di capi di una popolazione di animali selvatici. Essa consiste nel liberare fra la popolazione autoctona un numero noto di capi marcati in modo da poterli riconoscere. È intuibile che in

FASI IN CUI SONO STATI COMMESSI GLI ERRORI	NUMERO DI ERRORI COMMESSI	FASI IN CUI SONO STATI INDIVIDUATI GLI ERRORI				
		SPECIFICHE DEL SOFTWARE	PROGETTAZIONE DEL SOFTWARE	CODIFICA E PROVA DEL SOFTWARE	VALIDAZIONE DEL SISTEMA	ERRORI NON INDIVIDUATI
REQUISITI DEL SISTEMA	43	35	2	0	3	3
SPECIFICHE DEL SOFTWARE	42	28	7	4	3	0
PROGETTAZIONE DEL SOFTWARE	26		22	3	1	0
CODIFICA E PROVA DEL SOFTWARE	87			84	3	0

Gli sforzi compiuti per razionalizzare il processo di sviluppo del software, anche se riescono a ridurre il numero degli errori introdotti durante la progettazione e la codifica, non garantiscono la produzione di un software privo di errori. L'illustrazione mostra, nel caso di un'applicazione concreta, il numero degli errori commessi durante ciascuna fase del ciclo di sviluppo e inoltre la fase in cui questi sono stati individuati. I dati si riferiscono al Project on Diverse Software (PODS), un progetto sperimentale facente parte del programma di ricerca del gruppo Halden Reactor Project norvegese, che aveva lo scopo di valutare l'effetto di alcune metodologie sulla produzione del software. Durante la stesura delle 3500 linee di codice, richieste per il progetto, si sono commessi in totale 198 errori. Dall'illustrazione si può notare che la gran parte di essi è stata individuata nella stessa fase in cui gli errori erano commessi, probabilmente durante l'attività di verifica. Alla fine di tutte le attività di sviluppo sono state eseguite ulteriori prove (oltre 600 000 prove simulanti l'applicazione reale) i cui risultati sono stati confrontati con quelli di due programmi scritti in modo diverso, ma svolgenti le medesime funzioni. In questo modo si sono individuati tre ulteriori errori che erano sfuggiti a tutte le attività di sviluppo.

un'ipotetica campagna di caccia la percentuale di capi marcati catturati, rispetto al numero di quelli liberati, sarà la stessa di quella dei capi autoctoni catturati rispetto al loro numero totale che si voleva stimare. Analogamente in un programma viene inserito di proposito un certo numero di errori che si vanno ad aggiungere a quelli «autoctoni» involontari dovuti alla progettazione e quindi, in base a quelli individuati durante la fase di prova, si tenta di stimare il numero degli errori «autoctoni» residui.

Attualmente sono in corso molti progetti di ricerca che hanno l'obiettivo sia di migliorare il processo di sviluppo presentato, sia di integrarlo con nuove tecniche nelle sue fasi più critiche. In particolare alcune ricerche mirano a formalizzare il progetto sin dalla definizione del problema in quanto si è notato che, nonostante l'uso di metodologie di sviluppo, i requisiti sono possibili sedi di contraddizioni e di sottintesi che generano fraintendimenti nel passaggio alla fase successiva.

Sono stati così ideati dei metodi di sviluppo formali (quale il *Vienna Development Method*) che permettono una descrizione del sistema tale da poter essere sottoposta ad analisi matematica per individuare eventuali contraddizioni sin dalla fase dei requisiti. Inoltre la descrizione può essere analizzata automaticamente con un rigore simile a quello delle verifiche di correttezza sintattica, o, talvolta, essere animata per mostrare vari aspetti del comportamento del sistema permettendo così di individuare eventuali carenze rispetto ai requisiti.

Purtroppo i metodi formali sono complessi e costosi da applicare e i mezzi automatizzati che ne facilitano l'utilizzazione sono ancora in fase di sviluppo.

Altre ricerche cercano di identificare una serie di misure oggettive che rappresentino validi indicatori per prevedere la correttezza del prodotto finale. Tali misure dovrebbero essere utilizzabili per tenere sotto controllo e ottimizzare tutto il ciclo di sviluppo del sistema.

Un buon esempio sono le misure di complessità che possono essere utilizzate a partire dalla definizione del sistema fino alla stesura del codice. Nel primo caso aiutano nel decidere fino a che livello deve spingersi il processo di decomposizione della funzione da svolgere in sottofunzioni; nel secondo indicano se il singolo programma sia stato realizzato dal programmatore in modo sufficientemente semplice e lineare o se debbano essere fatti ulteriori sforzi in questa direzione. Purtroppo, anche se è intuibile che quanto più un programma è semplice tanto meno esso è soggetto a errori, non si è ancora riusciti a esprimere questo legame tramite un modello matematico che metta in relazione fra loro complessità e correttezza: attualmente da queste misure si riescono a ricavare solamente indicazioni di tipo qualitativo.

La razionalizzazione del processo di sviluppo del software ha prodotto metodi e tecniche che sono andati progressivamente consolidandosi, tant'è vero che la relativa disciplina ha preso il nome di «ingegneria del software» per richiamare i razionali metodi ingegneristici utilizzati usualmente nella progettazione dei sistemi industriali. Questa disciplina ha contribuito a individuare gli aspetti più critici dello sviluppo del software e le tecniche che allo stato attuale dell'arte meglio permettono di ridurre gli errori. Tuttavia nel caso di un prodotto software non banale poco si può dire sul grado di correttezza raggiunto.

A causa di questa incertezza, la progettazione di sistemi che usano calcolatori in applicazioni a rischio rilevante deve essere condotta tenendo in conto un ulteriore requisito: il sistema non deve creare condizioni potenzialmente pericolose, neanche in presenza di occasionali situazioni anomale. In particolare il software dovrà essere realizzato in modo da riconoscere e fronteggiare gli effetti dovuti a eventuali errori residui (tali effetti sono detti «malfunzionamenti»).

Per quanto riguarda il riconoscimento, sono state sviluppate diverse tecniche, ognuna delle quali orientata a individuare una specifica classe di malfunzionamenti. Consideriamo l'esempio di un calcolatore utilizzato nel controllo del traffico aereo. Elaborando i dati ricevuti da un radar che spazza l'intero arco dell'orizzonte fra 0 e 360 gradi, esso determina la rotta di avvicinamento di un aereo. È abbastanza ovvio che, se il risultato dell'elaborazione è di «aereo in avvicinamento da 370 gradi», il programma contiene un errore. Questo tipo di malfunzionamento può essere individuato inserendo un'ulteriore parte di programma che, prima di fornire il risultato elaborato, ne verifichi la plausibilità, comparandolo con intervalli di valori ammissibili.

Un'altra classe tipica di malfunzionamenti è quella caratterizzata da tempi di risposta del sistema inadeguati. A scopo esemplificativo si può ricordare la situazione di stallo dell'esempio relativo al «semaforo intelligente». Questa potrebbe essere riconosciuta controllando che il tempo impiegato dall'elaborazione in corso non superi un valore atteso prefissato.

Per quanto riguarda il modo di fronteggiare i malfunzionamenti, sono possibili diverse strategie. La più semplice consiste nell'avvisare l'operatore che qualcosa, nel calcolatore, non sta funzionando a dovere e nel richiedere un suo intervento per evitare situazioni di pericolo. Ovviamente questa strada è percorribile solo in particolari applicazioni, quali quella del monitoraggio delle condizioni cliniche di pazienti in area di terapia intensiva ove può essere richiesto l'intervento dell'infermiere tramite un allarme.

Nel caso di applicazioni diverse, il sistema potrebbe non essere presidiato, oppure il compito che esso svolge potrebbe richiedere risposte talmente veloci da non lasciare tempo sufficiente per un intervento umano. Si ricorre in questo caso a strategie differenti: si può far interrompere al calcolatore le elaborazioni in corso, per riprenderle dall'inizio, nella speranza che la configurazione degli ingressi, probabilmente mutata nel frattempo, non porti allo stesso tipo di malfunzionamento. Alternativamente si può optare per il cosiddetto «fallimento sicuro»: il calcolatore, al riconoscimento di un grave malfunzionamento, viene messo fuori servizio facendogli fornire un ultimo comando tale da evitare le potenziali situazioni di pericolo.

Facendo sempre riferimento all'esempio del «semaforo intelligente», la prima di queste due strategie consiste nell'interrompere la situazione di «stallo», iniziando un nuovo conteggio, nella speranza che il numero delle macchine in attesa nelle due direttrici non sia più uguale; la seconda strategia consiste nel mettere fuori servizio il microprocessore dopo avergli fatto portare il semaforo su «giallo lampeggiante» in tutte le direzioni.

Una tecnica completamente diversa per riconoscere e fronteggiare i malfunzionamenti dovuti a errori residui si basa sull'uso di software ridondante: la stessa funzione viene svolta da due o più programmi che risiedono generalmente su calcolatori diversi. Se i programmi sono identici essi dovranno basare le loro elaborazioni su ingressi forniti da strumenti di misura differenti, in modo che la configurazione degli ingressi dei due programmi non sia mai contemporaneamente la stessa e quindi sia ridotta la probabilità di malfunzionamenti comuni a entrambi.

Una soluzione ancora migliore consiste nell'usare, per lo stesso scopo, programmi diversificati, cioè basati su algoritmi di calcolo e strutture interne differenti. Per esempio l'area di un cerchio si potrebbe ottenere calcolando il prodotto del quadrato del raggio per *pi* greco, in un programma, e l'area del poligono regolare inscritto, al tendere del numero dei lati all'infinito, nell'altro. Se i due risultati saranno uguali, a meno di un predeterminato margine dovuto alle approssimazioni del calcolo, la soluzione sarà corretta. Altrimenti si potrà dire che vi è stato un malfunzionamento in uno dei due programmi (anche se non si sarà in condizione di dire in quale). Potrà perciò essere stampato un messaggio per l'operatore o intrapresa un'operazione correttiva. Nel caso in cui i programmi diversificati che svolgono la medesima funzione siano tre, allora potrà essere utilizzata la risposta fornita dalla maggioranza di essi (logica di votazione due su tre), ottenendo così un sistema che continua a funzionare regolarmente anche in presenza di un malfunzionamento.

Tutte queste soluzioni tecnologiche permettono di minimizzare i malfunzionamenti dei sistemi basati sui calcolatori nell'ambito di applicazioni a rischio rilevante. Accettando il principio secondo il quale nessuna attività può essere considerata perfettamente sicura, il problema diviene quello di basare l'introduzione di tali sistemi su una valutazione costi-benefici che tenga conto dei loro limiti, ma anche degli innegabili vantaggi che essi possono offrire, specialmente nell'ambito di specifiche applicazioni.

BIBLIOGRAFIA

RAMAMOORTHY C. V. e BASTANI F. B., *Software Reliability - Status and Perspectives* in «IEEE Transactions on Software Engineering», SE-8, n. 4, luglio 1982.

SHOOMAN MARTIN L., *Software Engineering*, McGraw-Hill, New York, 1983.

BISHOP P. e altri, *STEM - A Project on Software Test and Evaluation Methods*, Proceedings of the Safety and Reliability Society Symposium - Manchester, Elsevier Applied Science, 1987.

EUROPEAN WORKSHOP ON INDUSTRIAL COMPUTER SYSTEMS T.C.7, *Dependability of Critical Computer Systems 1*, Elsevier Applied Science, 1988.

La sintesi di elementi superpesanti

Recentemente, grazie all'impiego di nuove tecniche di fusione dei nuclei, sono stati prodotti alcuni isotopi che sono molto più pesanti degli elementi esistenti in natura e che vengono resi stabili da effetti quantomeccanici

di Peter Armbruster e Gottfried Münzenberg

Negli ultimi 20 anni i fisici di tutto il mondo sono stati impegnati nella produzione di elementi superpesanti. Presso la Gesellschaft für Schwerionenforschung (GSI, società per la ricerca sugli ioni pesanti) di Darmstadt abbiamo ottenuto alcuni successi essendo riusciti a sintetizzare i nuclei degli elementi caratterizzati da numeri atomici 107, 108 e 109. Questi nuclei si trovano al di là della soglia di 106 protoni che segnava il limite delle precedenti tecniche per la sintesi e l'identificazione di elementi pesanti.

Le misure sperimentali di massa, sorrette da considerazioni teoriche, indicano che i nuovi elementi devono la loro stabilità agli effetti quantistici prodotti dalla disposizione dei loro protoni e neutroni. D'altra parte abbiamo incontrato problemi che ci hanno finora impedito di raggiungere gli obiettivi che ci eravamo prefissati alla fine degli anni sessanta, quando sembrava imminente la sintesi di elementi fino al numero atomico 114. Le ricerche compiute per superare questi ostacoli, tuttavia, ci hanno permesso di approfondire le nostre conoscenze sulla struttura nucleare e sulla dinamica delle reazioni di fusione tra nuclei.

La nucleosintesi ha fatto molta strada dai suoi inizi, quando gli elementi non reperibili in natura venivano prodotti nei reattori nucleari. I fisici hanno usato «proiettili» sempre più pesanti per bombardare gli atomi bersaglio; il risultato più recente è la fusione fredda, nella quale le masse e le energie di bombardamento sono scelte accuratamente in modo da minimizzare l'eccitazione dei nuclei di nuova formazione.

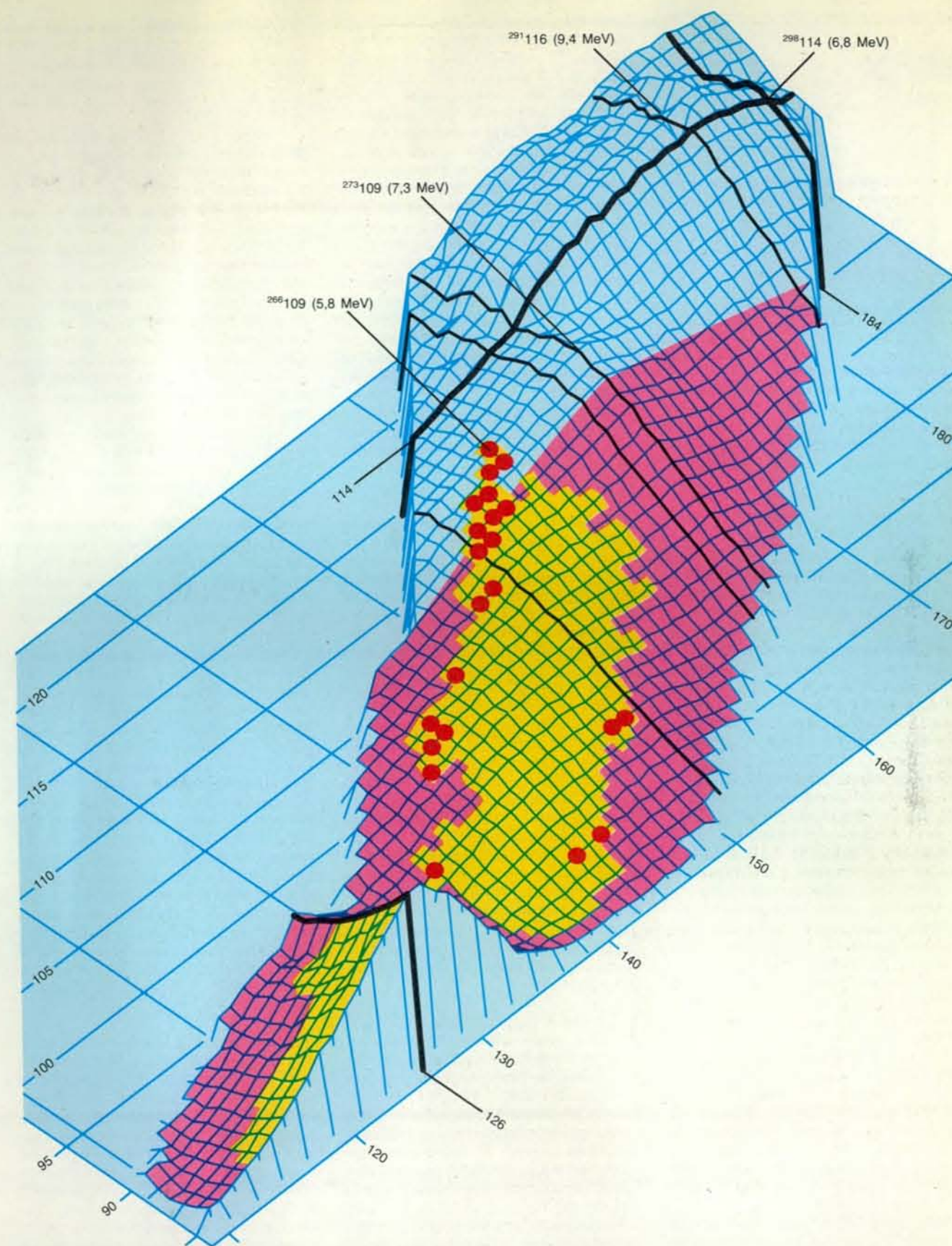
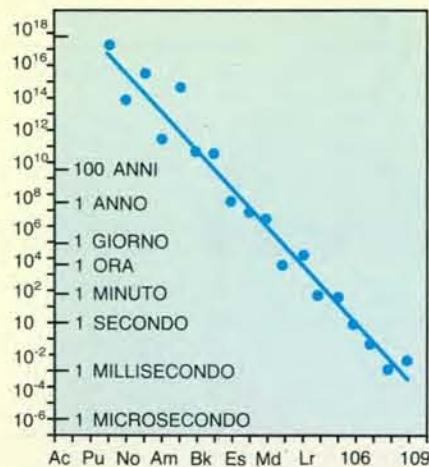
Nel corso delle nostre ricerche le idee iniziali sulla sintesi di elementi superpesanti si sono dimostrate quasi tutte errate: i nuclei degli elementi che si possono sintetizzare sono deformati e non sferici, come si credeva nel 1966. Nel processo di fusione utilizziamo nuclei stabili e sferici (abbondanti in natura) e proiettili di medio peso e non, come si riteneva in precedenza, i nuclei radioattivi artificiali più pesanti insieme a proiettili il più possibile leggeri. La fusione deve avvenire all'energia di bombardamento più bassa possibile, contrariamente alle prime opinioni secondo cui l'energia di impatto doveva essere sovrabbondante per favorire «violentemente» il processo mediante eccedenza di forze.

Nel 1934 Enrico Fermi bombardò atomi di tallio con neutroni lenti e produsse piombo mediante decadimento beta. Poiché per decadimento beta un neutrone dà luogo a un protone e a un elettrone, la sperimentazione in atto in quegli anni con bombardamenti neutronici fece scaturire l'idea di poter sintetizzare elementi transuranici (cioè con numero atomico maggiore di 92). I fisici si erano resi conto infatti che la cattura di neutroni seguita dal decadimento beta produce elementi con numero atomico maggiore di un'unità rispetto all'elemento originario.

Tra il 1940 e la metà degli anni cinquanta furono prodotti per bombardamento neutronico gli elementi 93, 94, 99 e 100. Per coincidenza, il fermio 100 fu l'ultimo della serie di elementi che si riuscì a produrre con il processo di cattura neutronica e decadimento beta ideato da Fermi: nessuno dei suoi isotopi presenta infatti decadimento beta. Nello stesso periodo vennero prodotti gli elementi dal 95 al 98 e il 101 per bombardamento alfa, un processo a seguito del quale nuclei pesanti assorbono due protoni e due neutroni, facendo in questo modo aumentare di due unità il numero atomico. Come tutti gli elementi pesanti, i transuranici contengono più neutroni che protoni. Per esempio, l'isotopo più comune del plutonio (elemento 94) presente in natura contiene 145 neutroni e ha numero di massa 239; l'isotopo del fermio (elemento 100) a vita media più lunga ha 157 neutroni e numero di massa 257.

Per produrre elementi con numero atomico superiore a 100, il metodo ovvio era quello di fondere i nuclei degli elementi più pesanti con nuclei di elementi leggeri contenenti più protoni e neutroni dell'elio. Gli elementi con numero atomico fino a 99 erano adatti allo scopo perché si potevano produrre in quantità macroscopiche pesabili. A Berkeley negli Stati Uniti e a Dubna in Unione Sovietica vennero costruiti acceleratori in grado di conferire agli ioni pesanti un'energia sufficiente a vincere le forze elettrostatiche che ostacolano la fusione. Tra il 1958 e il 1974 questi acceleratori di ioni pesanti consentirono la sintesi degli elementi dal 102 al 106. È ancora oggetto di disputa la priorità della scoperta di tali elementi, e quindi il privilegio di dar loro un nome.

Le metodologie che hanno funzionato così bene a Berkeley e a Dubna non hanno però avuto successo nella produzione di elementi con numero atomico superiore a 106. Per capire perché è così difficile sintetizzare gli elementi superpesanti - e perché alcuni di essi possono essere particolarmente stabili - è necessario conoscere in generale quali sono i meccanismi che mantengono compatti i nuclei o li frantumano e come cambia con l'aumentare del numero di massa l'equilibrio delle diverse forze che regolano la loro stabilità. Effetti trascurabili in nuclei più piccoli risultano essenziali nel determinare la differenza tra totale instabilità e vita media relativamente lunga nei nuclei più grandi.

La caratteristica essenziale di tutti i nuclei è l'interazione tra la forza nucleare forte, che tiene uniti protoni e neutro-

I gusci nucleari completamente occupati (*curve in nero*) caratterizzano i nuclei in cui il numero di protoni e neutroni contribuisce a una superstabilità. L'effetto dei gusci aumenta l'energia di legame di un nucleo e permette l'esistenza di isotopi che non sarebbero stabili se formati semplicemente da goccioline di «fluido nucleare». La barriera energetica di fissione per le goccioline nucleari (*in rosa*) scompare in corrispondenza dell'elemento 107. Gli effetti di guscio sono massimi nei pressi del piombo (*in basso a sinistra*) e aumentano di nuovo nella regione degli elementi superpesanti. I cerchietti indicano gli isotopi sintetizzati dal gruppo degli autori.

Il periodo di dimezzamento e con esso la vita media degli elementi radioattivi diminuisce all'aumentare della massa. Gli elementi più pesanti decadono quasi subito dopo la loro sintesi; di conseguenza sono necessari nuovi metodi per rivelarli e identificarli.

ni, e la forza elettrostatica, che tende ad allontanare i protoni l'uno dall'altro. Via via che il numero di massa aumenta, i nuclei diventano relativamente più ricchi di neutroni, il che ostacola le forze repulsive tra i protoni. Ciononostante, il legame tra i nucleoni presenta un massimo di intensità in corrispondenza del ferro (26 protoni e 30 neutroni), a meno di un quarto della tavola periodica degli elementi, per poi diminuire piuttosto rapidamente.

Benché la scissione di qualsiasi nucleo più pesante del ferro possa liberare energia, per frantumare nuclei più leggeri del piombo è necessaria un'energia tanto elevata che la reazione avviene solo in condizioni eccezionali. I nuclei più pesanti del piombo sono instabili, dato che possono raggiungere uno stato più fortemente legato espellendo una piccola parte del nucleo. Gli isotopi naturali dell'uranio e del torio decadono in genere emettendo una particella alfa. Soltanto nell'uranio e negli elementi più pesanti, nuclei non eccitati possono subire fissione spontanea.

In generale i nuclei atomici diventano sempre più instabili con il crescere del numero atomico: la loro vita media diminuisce da alcune migliaia di anni a pochi milionesimi di secondo. Tuttavia, le teorie della struttura nucleare ci inducono a ritenere che elementi di poco più pesanti di quelli finora sintetizzati possano essere più stabili.

Nuclei con determinate combinazioni di neutroni e protoni hanno energie di legame particolarmente alte: l'elio 4, l'ossigeno 16, il calcio 40, il calcio 48 e il piombo 208 sono molto stabili in confronto ai loro vicini nella tavola periodica. Queste elevate energie di legame sono dovute a una struttura a guscio che è l'equivalente nucleare dei gusci nei quali gli elettroni si distribuiscono attorno al nucleo. Le configurazioni di nucleoni che formano gusci completamente occupati risultano particolarmente stabili. Nel caso del piombo per esempio, la struttura a guscio fa aumentare l'energia di legame del nucleo di 11 milioni di elettronvolt (MeV) rispetto a quella di un'ipotetica «gocciolina» con lo stesso numero di protoni e di neutroni ma priva di struttura. Per la maggior parte dei nuclei, che hanno energie di legame fino a due miliardi di elettronvolt, questo aumento è relativamente insignificante. Tuttavia per gli elementi più pesanti dotati di una stabilità ridotta, la stabilizzazione dovuta alla presenza dei gusci può rappresentare la differenza tra una disintegrazione istantanea e una vita relativamente lunga.

I nuclei con gusci completamente occupati sono, come abbiamo detto, particolarmente stabili; il primo elemento che può avere un guscio completo dopo il piombo deve possedere 114 protoni e 184 neutroni. Il successo del modello a gusci nel prevedere le energie di legame dei nuclei più leggeri ha fatto sorgere l'ipotesi che i nuclei con peso prossimo a 298 possano essere tanto stabilizzati da formare - come nel caso dell'uranio e del torio - un'«isola» di elementi relativamente stabili. Tuttavia, questi elementi superpesanti stabilizzati, al contrario di quelli della regione uranio-torio, risulterebbero instabili se fossero goccioline omogenee di materia nucleare.

Il primo degli elementi superpesanti stabilizzati, quello con numero atomico 107, le cui proprietà coincidono con quelle dell'ekarenio postulato da Fermi, è stato identificato a Darmstadt nel 1981, 47 anni dopo che ne era stata ipotizzata l'esistenza. Da tale data abbiamo prodotto e identificato anche gli elementi con numeri atomici 108 e 109. La misurazione delle loro energie di legame dimostra che siamo già arrivati nella regione degli elementi superpesanti; stiamo ora studiando quali vincoli si applichino alla produzione di elementi ancora più pesanti.

La sintesi di elementi pesanti per fusione richiede allo sperimentatore di stabilire un delicato equilibrio tra i metodi di bombardamento che non producono alcuna reazione e quelli che provocano la fissione del nucleo che si è ottenuto, anziché portarlo in uno stato relativamente stabile. Il ridotto riscaldamento del nucleo di nuova formazione è il motivo principale per abbandonare la tecnica del bombardamento di nuclei bersaglio relativamente pesanti con ioni leggeri in favore del bombardamento di bersagli meno pesanti con ioni relativamente più pesanti (una variante eseguita per la prima volta a Dubna da Yuri Oganessian e colleghi). Per esempio, quando piombo 208 o bismuto 209 vengono fusi con cromo 54 o con ferro 58, l'energia di eccitazione del nuovo nucleo è di circa 20 MeV. Invece nella fusione di attinidi pesanti (californio 249, berkelio 249 o curio 248) con carbonio 12, azoto 15 od ossigeno 18 si ha un'energia di eccitazione di circa 45 MeV.

Il nuovo nucleo formato dalle fusioni di ioni leggeri e di attinidi raggiunge uno stato stabile mediante emissione di quattro neutroni, mentre quello formato da piombo o bismuto e da ioni pesanti emette un solo neutrone. Poiché la probabilità che un nucleo si raffreddi emettendo un neutrone è molto bassa rispetto alla probabilità di subire fissione, ogni fase di emissione di un neutrone subita da un nucleo riduce in modo significativo la resa finale in nuclei superpesanti. Con il meccanismo di emissione di un solo neutrone si ha una probabilità molto maggiore che un nucleo di nuova formazione rimanga intatto.

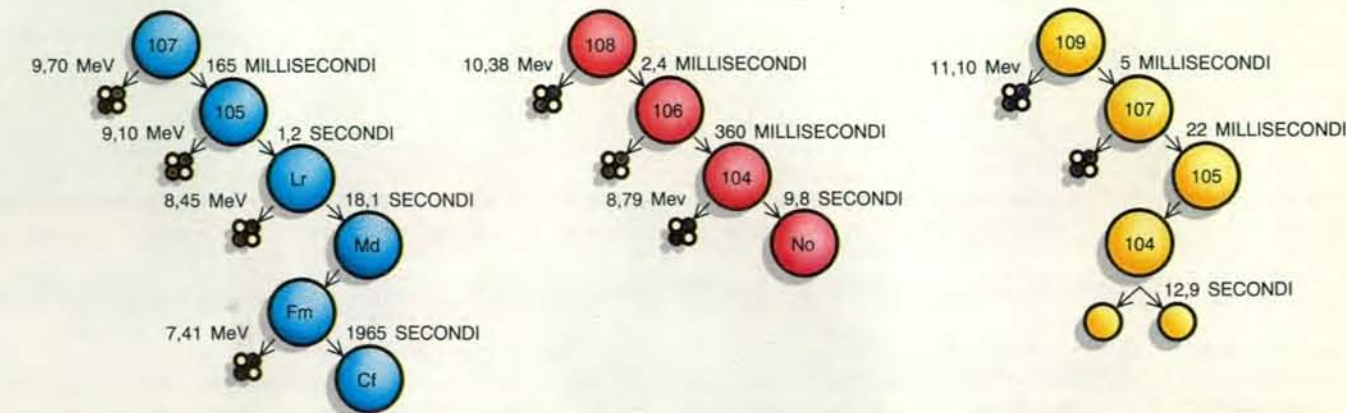
Purtroppo la fusione fredda presenta anche un inconveniente: le forze elettrostatiche repulsive tra i due nuclei rendono più difficile la loro fusione. Quando due nuclei si avvicinano uno all'altro, una parte della loro energia cinetica si trasforma in energia di eccitazione del sistema durante la collisione e non è più disponibile per superare la barriera energetica di fusione, facendo in tal modo diminuire la probabilità che i nuclei si fondano.

Nella fusione fredda con ioni pesanti si ha la conversione di una quantità maggiore di energia cinetica durante la scalata e l'attraversamento della barriera energetica e quindi la probabilità di superare quest'ultima viene molto ridotta in confronto a una reazione tra ioni leggeri e bersagli molto pesanti. Se si aumenta l'energia iniziale per compensare tali perdite, l'energia di eccitazione cresce e diminuisce il numero di nuclei che rimangono intatti; ne consegue che la fusione fredda è vantaggiosa solo per nuclei con numeri atomici superiori a 106. Con le nostre ricerche abbiamo dimostrato che le sezioni d'urto più grandi per la formazione di elementi pesanti si trovano in un ristretto intervallo di energia di circa 5 MeV al di sopra della barriera energetica di fusione.

Le considerazioni teoriche sulla sintesi di nuclei superpesanti possono costituire un interessante esercizio, ma la pratica è un affare molto più complesso. Le teorie devono misurarsi con gli aspetti progettuali degli acceleratori e dei bersagli e con lo sviluppo di rivelatori che siano in grado di identificare un nucleo superpesante una volta che questo sia stato sintetizzato. Quando l'idea di produrre nuclei superpesanti stimolò per la prima volta l'immaginazione di fisici e chimici alla fine degli anni sessanta, nessuno in Germania era esperto di nucleo-sintesi. Per i nuovi adepti del campo le strade aperte erano diverse; c'era molto da imparare dai vecchi esperimenti di Dubna e di Berkeley, ma era chiaro che non si sarebbero fatti ulteriori progressi copiandoli. Era necessario disporre di un acceleratore di ioni pesanti, di metodi di separazione veloci per isolare i nuovi elementi e di tecniche adeguate per la loro identificazione. E nessuno era ancora in grado di prevedere esattamente quali reazioni nucleari avrebbero avuto successo.

Nel 1969 il Governo federale e il Land dell'Assia raggiunsero un accordo per finanziare congiuntamente un nuovo ente di ricerca sugli ioni pesanti (il GSI appunto) con sede a Darmstadt. L'Acceleratore lineare universale (UNILAC), che rappresenta il fulcro dell'attività della GSI, è entrato in funzione nel 1975.

L'UNILAC è in grado di accelerare tutti gli ioni, fino all'uranio compreso, a energie superiori a quella della barriera coulombiana. La macchina è stata progettata con lo scopo specifico di produrre le correnti ioniche che fossero le più elevate possibili; una cura particolare è stata dedicata alla possibilità di variare di piccole quantità l'energia degli ioni e di poterla fissare riproducibilmente a un dato valore. Il progetto dell'acceleratore, che inizialmente venne elaborato da Christoph Schmelzer e colleghi di Heidelberg, prende spunto anche da precedenti ricerche effettuate altrove: le sorgenti di ioni sono una versione modificata di quelle usate a Dubna per produrre ioni di carica elevata e la struttura di Alvarez sviluppata a Berkeley è stata incorporata nell'acceleratore lineare ad alta frequenza.

Mentre l'UNILAC era ancora in fase di costruzione, molti gruppi di ricerca si posero il problema di come utilizzare al meglio le possibilità della macchina: quali reazioni e quali tecniche si dovevano impiegare? In un primo tempo l'UNILAC venne usato per verificare un gran numero di ipotesi, ma alla fine prevalse una sola strategia: la fusione fredda accoppiata al trasporto per rinculo dei prodotti di fusione.

Dalla scoperta del plutonio, nel 1941, sono state sintetizzate circa 400 tonnellate di tale elemento, una quantità corrispondente a 10^{30} atomi. D'altro canto sono stati prodotti e identificati solo pochi atomi dell'elemento 109. Perché gli elementi più pesanti sono stati prodotti in quantità così insignificanti? La risposta è che, per produrre plutonio, blocchi di uranio 238 con uno spessore dell'ordine di diversi centimetri sono stati bombardati con un numero enorme di di neutroni; invece l'UNILAC ha accelerato soltanto 100 microgrammi di ferro 58 su bersagli di piombo 208 dello spessore di poche centinaia di nanometri. Inoltre, la sezione d'urto di cattura di neutroni, la reazione che produce il plutonio 239, è circa 10 bilioni di volte superiore alla sezione d'urto di fusione nucleare, che produce l'elemento 109.

La crescente difficoltà di produrre elementi sempre più pesanti è solo una faccia della medaglia. Una volta sintetizzati, elementi quale il 109 decadono così rapidamente che la sintesi non può reggere il passo del decadimento. Gli elementi più pesanti hanno una vita media tanto breve che al termine del bombardamento tutti gli atomi che sono stati prodotti hanno già subito decadimento; questi atomi devono perciò essere rivelati e identificati durante il processo di produzione stesso.

Per elementi fino al 106 le tecniche di produzione e di rilevamento si basavano essenzialmente su metodi meccanici per trasportare gli elementi appena prodotti dalla regione di reazione fino a un rivelatore. Il tempo di trasporto tra la sintesi e il rilevamento dei prodotti di reazione era determinato dalla velocità di flusso dei prodotti di reazione in un gas, dal loro tempo di diffusione a partire da una superficie solida o dalla velocità di un bersaglio circolare rotante. Queste tecniche non erano però sufficientemente precise per rivelare elementi più pesanti del 106. Esse imponevano inaccettabili compromessi tra velocità e precisione del rilevamento, tanto che i metodi più veloci non erano in grado di identificare con certezza gli isotopi che erano stati appena sintetizzati.

La tecnica da noi adottata a Darmstadt è stata quella di sfruttare la velocità di rinculo impartita ai prodotti di reazione dagli ioni pesanti incidenti per trasportare al rivelatore i nuclei sintetizzati. Quando uno ione pesante colpisce un atomo del bersaglio e si fonde con esso, il nucleo risultante prosegue lungo la traiettoria originaria dello ione a una velocità che può arrivare fino a una frazione considerevole della velocità della luce. Il risultato è che si possono rivelare nuclei con un periodo di dimezzamento di soli 100 nanosecondi.

Anche se la tecnica di trasporto per rinculo permette di rivelare e identificare nuclei con vita media brevissima, essa rende però più complessa l'apparecchiatura di rilevamento. Dalla regione di reazione non escono solo i singoli nuclei veloci prodotti nella regione di fusione, ma anche bilioni di ioni pesanti e migliaia di atomi espulsi dalla lamina del bersaglio. Per separare i nuclei superpesanti dal resto del fascio, abbiamo installato un filtro di velocità (SHIP, da Separator for Heavy-Ion Reaction Products), messo a punto in collaborazione con il Secondo istituto di fisica dell'Università di Giessen. Per la cinematica della collisione e della reazione di fusione, tutti i prodotti di fusione rinculano a una velocità che può essere calcolata preventivamente; essi si possono quindi separare in modo relativamente semplice.

Il filtro è costituito da due stadi, ciascuno dei quali impiega campi elettrici e magnetici. I due campi tendono a deflettere in direzioni opposte oggetti carichi; soltanto se un nucleo è in moto alla velocità corretta gli effetti si annullano, ed esso prosegue lungo il piano mediano dell'apparecchiatura. Il nostro filtro tandem riduce il numero di ioni accelerato

Gli elementi superpesanti 107, 108 e 109 sono stati identificati in base a catene di decadimento osservate presso la Gesellschaft für Schwerionenforschung (GSI) di Darmstadt. I rivelatori registravano la massa del prodotto di fusione incidente, poi una serie di emissioni di particelle alfa e, nel caso dell'elemento 104, fissione spontanea. La probabilità che tali segnali correlati siano dovuti al caso è inferiore a una parte su 1000 bilioni; una sola serie di decadimenti può quindi dimostrare l'esistenza di un elemento superpesante.

che entrano nella regione di rilevamento di un fattore pari a 100 miliardi e il numero di nuclei di rinculo provenienti dal bersaglio di un fattore 1000.

Il filtro di velocità, pur eliminando dal fascio quasi tutte le particelle indesiderate, lascia passare più del 40 per cento dei prodotti di fusione. I rivelatori posti a valle dell'apparecchio registrano le catene di decadimento delle particelle che escono dal filtro e ci consentono di identificare con assoluta certezza i prodotti di fusione.

Il primo elemento del sistema finale di rilevamento è un dispositivo a tempo di volo che misura per la terza volta la velocità di una particella (le prime due misurazioni sono quelle su cui si basa il funzionamento dei filtri di velocità). Una volta attraversato il rivelatore a tempo di volo, la particella incide su rivelatori «spaziali» di silicio a barriera superficiale che ne registrano la posizione e l'energia. La combinazione di tempo di volo e energia consente di determinare approssimativamente la massa delle particelle e quindi di distinguere i prodotti di fusione dagli ioni diffusi e dai nuclei del bersaglio che rinculano.

Per identificare con certezza un nucleo è necessario però che il suo decadimento sia correlato a quello dei suoi prodotti in una serie di decadimento radioattivo. I processi di decadimento originati da un singolo nucleo hanno tutti le medesime coordinate spaziali e i nuclei che derivano da esso subiscono processi di decadimento il cui tipo, energia e periodo di dimezzamento sono noti da precedenti misurazioni.

Conoscendo tali decadimenti correlati, è possibile identificare in modo inequivocabile ogni nucleo prodotto nella fusione. Anche se una particella spuria situata nella stessa posizione dell'ipotetico prodotto di fusione potrebbe decadere e provocare un segnale spazialmente correlato, è estremamente improbabile che l'energia e il tipo del decadimento e il periodo di dimezzamento coincidano con quelli previsti per il prodotto di fusione. Abbiamo osservato queste catene di decadimento fino alla quarta generazione; la probabilità che una simile serie di eventi correlati si verifichi per caso è compresa tra uno su 10^{18} e uno su 10^{21}. Se eventi correlati prodotti dall'isotopo voluto si verificassero al tasso di uno al giorno, dovremmo aspettare per un tempo pari a 100 volte l'età della Terra per osservare un segnale spurio relativo a quattro generazioni di decadimenti. Ne consegue che anche un solo evento può dimostrare inequivocabilmente l'esistenza di un determinato isotopo superpesante.

Tra il 1981 e il 1986 noi, assieme ai nostri collaboratori Peter Hessberger, Sigurd Hofmann, Matti Leino, Willibord Reisdorf e Karl-Heinz Schmidt abbiamo utilizzato l'UNILAC e il suo sistema di rilevamento per sintetizzare e identificare gli elementi dal 107 al 109. Durante i nostri esperimenti abbiamo prodotto 14 isotopi di elementi dal numero atomico 104 al 109 (cinque dei quali già noti in precedenza) e due altri isotopi degli elementi con numero atomico 107 e 108, aventi rispettivamente numeri di massa 261 e 264.

Nel 1981 abbiamo prodotto un isotopo dell'elemento 107 con numero di massa 262 bombardando bismuto 209 con cromo 54. Per l'isotopo dispari-dispari dell'elemento 107 - nel quale cioè sia il numero di protoni, sia quello di neutroni sono dispari - conosciamo cinque energie di decadimento alfa, che forniscono informazioni sui livelli energetici del nucleo; possiamo affermare che l'isotopo possiede anche un isomero (uno stato eccitato a lunga vita).

L'elemento 109 è stato identificato in base a una singola catena di decadimento, osservata alle 16,10 del 29 agosto 1982 nella reazione tra ferro 58 e bismuto 209. Il nucleo $^{266}109$ è sopravvissuto per cinque millisecondi prima di emettere una particella alfa con energia di 11,1 MeV; il nucleo risultante dell'elemento 107 è decaduto dopo 22 millisecondi in un nucleo dell'elemento 105; quest'ultimo ha prodotto un nucleo 104 che dopo 12,9 secondi ha subito fissione spontanea. Da questa unica catena si sono potuti determinare, sia pure con limitata precisione, l'energia del decadimento, il periodo di dimezzamento e la sezione d'urto. Altre due catene di decadimento sono state osservate all'inizio del 1988, sei anni dopo l'identificazione dell'elemento 109; esse confermano l'interpretazione dell'evento osservato nel 1982.

Nel 1984 abbiamo identificato tre catene di decadimento dell'isotopo $^{265}108$ nella reazione che avviene tra ferro 58 e piombo 208. I due isotopi identificati degli elementi 107 e 109 sono isotopi dispari-dispari con probabilità di fissione assai bassa, mentre l'isotopo dell'elemento 108 ha un numero pari di protoni e dispari di neutroni. Sebbene la probabilità che un isotopo pari-dispari subisca fissione sia significativamente maggiore rispetto a un isotopo dispari-dispari, $^{265}108$ presenta decadimento alfa.

È particolarmente interessante il fatto che nessun isotopo degli elementi dal 107 al 109 subisca fissione spontanea e che gli isotopi pari-pari $^{256}104$, $^{260}106$ e $^{264}108$ presentino tutti una stabilità quasi identica nei confronti della fissione spontanea. Il livello di stabilità approssimativamente costante dimostra che gli effetti di stabilizzazione dovuti alla presenza di gusci bilanciano la diminuzione complessiva di stabilità che si verifica con l'aumentare della massa.

Al di là degli elementi 104 e 105 si trova dunque una piccola «isola» di nuclei che decadono per emissione alfa in isotopi noti di elementi più leggeri; questi decadimenti alfa ci permettono di determinare l'energia di legame degli elementi superpesanti. In ogni fase del decadimento, se si conosce l'energia di legame del nucleo derivato, è possibile calcolare quella del nucleo iniziale dalla sua energia di decadimento alfa. Per una catena di decadimenti alfa è possibile ricavare l'energia di legame del membro iniziale della serie se si conosce l'energia di legame del prodotto finale. La catena di decadimento dall'isotopo $^{264}108$ al $^{260}106$ al $^{256}104$ al $^{252}102$ può essere ricostruita perché è stato osservato un evento relativo al decadimento degli elementi 108 e 104 e per l'elemento 106 ne sono stati osservati diversi. Le energie di legame di questi nuclei sono rispettivamente 120, 106 e 94 MeV.

La correzione dell'energia di legame dovuta all'effetto dei gusci aumenta in modo regolare per tutti gli isotopi compresi fra l'uranio 232 e il $^{264}108$ che sono collegati da processi di decadimento alfa; i corrispondenti valori si innalzano, da 1-2 MeV fino a 6-7 MeV. In effetti, gli elementi che vanno dall'uranio all'elemento 108 hanno barriere energetiche di fissione ugualmente alte, pari a circa 6 MeV. Diversamente dall'uranio, che dovrebbe ancora essere stabile in forma di aggregato di nucleoni, gli elementi 106 e 108 devono la loro stabilità solo alla disposizione quantomeccanica dei loro sistemi di fermioni a molti corpi. Recenti ricerche teoriche prevedono barriere energetiche di fissione in accordo con le nostre misure.

Sono l'altezza e lo spessore di tale barriera a determinare in gran parte la vita media di un elemento nei confronti della fissione. L'effetto stabilizzante dei gusci fa aumentare la vita media degli elementi 106 e 108 di 15 ordini di grandezza. Su scala logaritmica le vite medie osservate vengono a cadere tra l'*eigenzeit* di un nucleo (il tempo di circa 10^{-21} secondi che sarebbe necessario perché un insieme di nucleoni non legati si disgreghi) e l'età dell'universo (circa 10^{18} secondi). È solo in confronto alla durata della vita umana (2×10^9 secondi) che i nuovi elementi sono instabili; per renderli stabili su tale scala la loro vita media dovrebbe essere di 12 ordini di grandezza maggiore. La fisica nucleare, però, non si basa su una scala dei tempi umana.

L'isola di attività alfa da noi scoperta è una conseguenza diretta della stabilizzazione di questi isotopi dovuta all'effetto dei gusci. Così, la stabilizzazione prevista alla fine degli anni sessanta per i nuclei sferici superpesanti attorno all'elemento 114 comincia in realtà molto prima di quanto si ritenesse e cresce costantemente. Nello stretto minimo di instabilità dopo il piombo, tra gli elementi con numero atomico 83 e 90, l'effetto stabilizzante dei gusci diminuisce, ma tra gli elementi con numero atomico 92 e 114 l'energia di stabilizzazione aumenta lentamente e regolarmente.

Anche i nuclei corrispondenti al «ponte» che collega l'isola di nuclei superpesanti con la «terraferma» del resto del sistema periodico sono stabilizzati dall'ordinamento quantistico del sistema di fermioni, mentre la maggior parte dei nuclei è stabilizzata dalle proprietà di massa delle gocce liquide. I nuclei degli elementi con numero atomico dal 107 al 109 si trovano sul ponte tra l'isola e la terraferma ed è solo una questione di punti di vista stabilire se i nuovi isotopi appartengono all'isola o alla terraferma. In ogni caso anch'essi - come gli elementi superpesanti - possono essere rilevati solo a causa della stabilizzazione dei loro stati fondamentali dovuta alla presenza di gusci.

Se consideriamo le previsioni più recenti della teoria della stabilizzazione dell'energia di legame, vediamo che dovrebbe esistere una regione di quasi 400 nuclei superpesanti tra gli elementi 106 e 126 la cui barriera energetica di fissione supera i 4 MeV. Tutti questi isotopi dovrebbero avere vita media superiore a un microsecondo; se venissero sintetizzati, dovrebbe essere possibile rilevarli con le tecniche esistenti. Regioni di particolare stabilità sono previste in prossimità degli isotopi $^{273}109$ e $^{291}115$. In corrispondenza di un numero di neutroni pari a circa 166, la deformazione dello stato fondamentale cambia: gli isotopi più leggeri di quest'ultimo sono deformati, mentre quelli più pesanti sono sferici.

Negli ultimi 20 anni tutti i tentativi di produrre isotopi ricadenti nel previsto centro di stabilizzazione intorno all'isotopo $^{298}114$ sono falliti. Né la fusione né altre reazioni con ioni pesanti hanno permesso di identificare questi isotopi superpesanti. Nondimeno, l'idea fondamentale che possano esistere sistemi di nucleoni stabilizzati dall'effetto dei gusci oltre la regione degli aggregati nucleari stabili è stata confermata dagli esperimenti qui descritti. Vi sono parecchi buoni motivi per credere che la teoria possa essere estrapolata a elementi ancora più pesanti.

Si pone ora un'interessante domanda:

Un filtro efficiente è fondamentale per il rilevamento e l'identificazione di elementi superpesanti; altrimenti i segnali provenienti dal fascio di ioni che bombarda il bersaglio potrebbero sovraccaricare i rivelatori. Il filtro SHIP (separatore degli ioni pesanti prodotti nella reazione) della GSI è formato da due gruppi di deflettori elettrostatici e magnetici che consentono il passaggio solo di ioni aventi la velocità corretta. Gli ioni che non presentano caratteristiche in accordo con i previsti prodotti di fusione vengono deviati al di fuori della traiettoria del fascio. Il filtro riduce il numero di ioni pesanti nel fascio di un fattore 100 miliardi e quello di nuclei rinculanti del bersaglio di un fattore 1000. Un nucleo superpesante invece ha una probabilità superiore a due su cinque di raggiungere il rivelatore.

Gli isotopi superpesanti sono caratterizzati dalla resistenza al decadimento per fissione (*in blu*). Invece di subire fissione essi decadono per emissione di particelle alfa (*in rosso*) o per cattura di elettroni (con la quale un protone si trasforma in un neutrone, riducendo di un'unità il numero atomico). Contrariamente alla tendenza di una ridotta stabilità con l'aumentare del numero atomico, il periodo di dimezzamento dell'elemento 109 è maggiore di quello del 108. La cornice in nero evidenzia gli isotopi sintetizzati dagli autori.

che cosa limita in ultima analisi la sintesi di questi «fragili oggetti»? Grazie ai nostri studi intensivi delle reazioni di fusione abbiamo scoperto alcuni dati importanti. I nuclei stabilizzati con stati fondamentali sferici possono essere distrutti eccitandoli a un'energia di soli 15 MeV - come è stato dimostrato sperimentalmente da Karl-Heinz Schmidt fin dal 1979 - mentre i nuclei deformati possono sopportare un'eccitazione fino a 40 MeV. Anche nella reazione tra calcio 48 e curio 248 - la più conveniente fra le reazioni accessibili - si trova un'energia di eccitazione di circa 30 MeV. Ciò significa che è possibile produrre solamente elementi superpesanti con nuclei deformati; finora questi tentativi hanno però avuto successo solo per elementi con numero atomico inferiore a 110.

La fusione di due nuclei con formazione di un nucleo superpesante è, come abbiamo detto sopra, ostacolata fin dall'inizio dall'esigenza di superare la barriera energetica di fusione. Per un dato nucleo prodotto, la probabilità di superare tale barriera è massima quando ioni il più possibile leggeri bombardano bersagli il più possibile pesanti. Nonostante il vantaggio di facilitare la fusione, questa combinazione estremamente asimmetrica ha lo svantaggio di produrre il massimo riscaldamento del nucleo prodotto e di causare quindi le massime perdite per fissione quando il nucleo ricade in uno stato meno eccitato. Per combinazioni meno asimmetriche le perdite nella fase di raffreddamento sono proporzionalmente minori. Il miglior compromesso tra le perdite nello stadio finale e la probabilità di formazione nello stadio iniziale è offerto dalle combinazioni più simmetriche con nuclei bersaglio nella regione del piombo.

L'impiego del piombo e del bismuto come bersagli trae un duplice vantaggio dalla stabilizzazione dovuta ai gusci in tali nuclei: il fatto che essi, con i loro gusci doppiamente completi, siano fortemente legati porta a una riduzione di più di 10 MeV dell'energia fornita al nucleo prodotto e a una corrispondente diminuzione delle perdite per fissione. Inoltre, la probabilità di superare la barriera energetica di fusione è maggiore quando si usano come partner nella reazione nuclei sferici, fortemente legati e relativamente duri. Anche qui i forti effetti di guscio del piombo fanno valere i loro diritti, questa volta, però, nella dinamica del processo.

Stiamo cominciando ora a capire perché sarà difficilissimo produrre elementi ancora più pesanti. Soltanto le interazioni fra la stabilizzazione dovuta ai gusci nei nuclei che partecipano alla fusione e hanno gusci completi, gli effetti di guscio nella dinamica della reazione di fusione e la maggiore stabilità dei nuclei superpesanti deformati allo stato eccitato ci permettono di sintetizzare alcuni isotopi dei più leggeri elementi superpesanti. Dobbiamo ampliare la domanda iniziale sull'esistenza dei nuclei stabilizzati all'effetto della presenza di gusci in tutti gli stadi del processo di sintesi. Il mantenimento di un ordine esistente, evitando di introdurre disordine nella reazione di fusione è essenziale per la produzione di questi oggetti complessi e fragili.

In che modo si potranno produrre altri elementi superpesanti? Per gli elementi 110 e 111 dovrebbe essere possibile ricorrere ai metodi che abbiamo messo a punto, con reazioni tra nichel 62 e piombo 208 o bismuto 209. Quando gli elementi saranno stati prodotti, la loro individuazione sarà una questione non tanto di nuove capacità tecniche, quanto di metodi di arricchimento isotopico associati alla pazienza di perfezionare e applicare la tecnica per mesi e mesi.

BIBLIOGRAFIA

MÜNZENBERG G. e altri, *Identification of Element 107 by α-Correlation Chains* in «Zeitschrift für Physik A», 300, n. 1, giugno 1981.

MÜNZENBERG G. e altri, *Evidence for Element 109 from One Correlated Decay Sequence Following the Fusion of ^{58}Fe with ^{209}Bi* in «Zeitschrift für Physik A», 315, n. 1, gennaio 1984.

ARMBRUSTER PETER, *On the Production of Heavy Elements by Cold Fusion: The Elements 106 to 109* in «Annual Review of Nuclear and Particle Science», 35, pp. 135-194, 1985.

MÜNZENBERG G. e altri, *Observation of the Isotopes ^{264}108 and ^{265}108* in «Zeitschrift für Physik A», 328, n. 1, settembre 1987.

MÜNZENBERG GOTTFRIED, *Recent Advances in the Discovery of Transuranium Elements* in «Reports on Progress in Physics», 51, n. 1, gennaio 1988.

La topobiologia

Questo ramo della biologia che si occupa delle interazioni dipendenti dalla localizzazione delle cellule, in particolare nell'embrione in via di sviluppo, consente di chiarire l'evoluzione del sistema immunitario

di Gerald M. Edelman

Nella scienza, uno degli aspetti più gratificanti è proprio l'impossibilità di prevedere dove condurrà esattamente un particolare indirizzo di ricerca. In effetti, un'indagine estesa può finire per fare luce su qualche vecchio enigma in un campo che sembrava inizialmente distante. Una connessione tanto inaspettata è emersa di recente tra due settori della biologia ai quali mi sono dedicato per la maggior parte della mia carriera scientifica. Il primo settore è quello che prende in considerazione la struttura degli anticorpi, cioè di quelle molecole del sistema immunitario che riconoscono gli agenti invasori e innescano i meccanismi che portano alla loro distruzione. Esso ha raggiunto il massimo dello sviluppo alla fine degli anni sessanta con la definizione della complessa struttura degli anticorpi.

In seguito, agli inizi degli anni settanta, ci si è resi conto che gli anticorpi appartengono, insieme ad altre molecole del sistema immunitario, a una singola entità evolutiva, la superfamiglia delle immunoglobuline. Si trattava di un risultato sorprendente, ma forse non lo accolsi con lo stesso entusiasmo con cui lo avrei accolto alcuni anni prima, dato che a quell'epoca la mia attenzione era rivolta ormai a un altro problema: come, in un embrione in via di sviluppo, le cellule interagiscono per dare origine a un nuovo organismo. Questo nuovo lavoro ha condotto alla scoperta di molecole che favoriscono l'adesione tra le cellule e che sono dette CAM (dall'inglese *cell-adhesion molecules*). Si tratta di proteine che regolano le interazioni tra le cellule dell'embrione. Analisi più recenti su di esse e su molecole affini hanno fornito nuovi elementi che hanno permesso di porre le basi di un'embriologia molecolare che leghi la forma e la funzione dei tessuti embrionali all'evoluzione e alla genetica. La maggior parte di questo articolo sarà dedicata alla descrizione di queste recenti indagini.

Ma prima di arrivare alla conclusione tornerò al tema del mio precedente lavoro. La ragione di questo ritorno è che la scoperta e l'analisi delle CAM hanno risolto, per ironia della sorte, l'annoso problema dell'origine evolutiva della superfamiglia delle immunoglobuline. Di recente è stato dimostrato che i geni che codificano per le CAM e quelli che codificano per le immunoglobuline hanno un così alto numero di sequenze simili nel DNA e nelle corrispondenti proteine che devono necessariamente essere correlati dal punto di vista evolutivo. Inoltre, le CAM sono diffuse in tutto il regno animale, mentre il sistema immunitario adattativo è limitato ai vertebrati. Sembra probabile, pertanto, che le molecole di quest'ultimo siano derivate dai geni che controllano il sistema dell'adesione tra cellule, fornendo un esempio di notevole opportunismo evolutivo.

Sia gli anticorpi sia le molecole che favoriscono l'adesione tra cellule esercitano i loro effetti principali in corrispondenza delle superfici cellulari. Le interazioni superficiali tra cellule possono condurre a cambiamenti nell'espressione genica e nella forma, nella motilità e nella funzione delle cellule. Quali reazioni avvengono esattamente durante l'interazione tra le cellule dipende in parte dalla storia stessa delle cellule, cioè da quali interazioni esse hanno avuto in passato con altre cellule. E, cosa ancora più importante, dato che queste interazioni dipendono dal tipo di cellule che circondano una data cellula, le reazioni saranno differenti a seconda della sede in cui si verificano. Queste interazioni sito-dipendenti si instaurano effettivamente nel sistema immunitario, ma rivestono una particolare importanza nello sviluppo embrionale. In effetti, il destino delle cellule (ciò che esse diventeranno nell'organismo maturo) dipende in modo cruciale dalla sede che quelle cellule occupano nell'embrione e da ciò che le circonda.

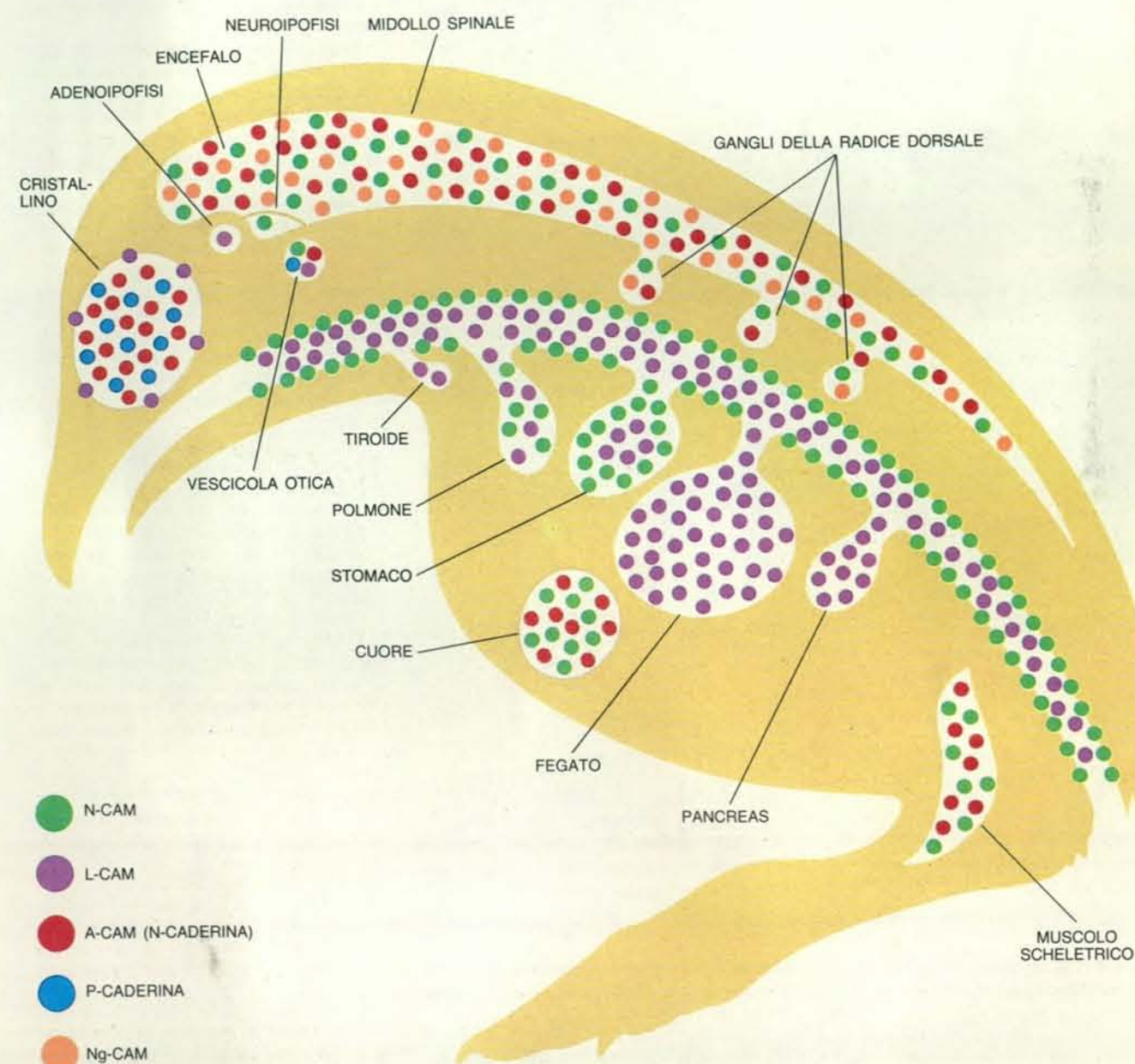
Benché da tempo gli embriologi fossero consapevoli, in generale, della possibilità di considerare la localizzazione un elemento critico nel determinare il destino cellulare, è sembrato indispensabile rimettere a fuoco il problema, dando un nome allo studio di quelle interazioni sito-dipendenti che si svolgono alla superficie delle cellule e che conducono alla regolazione cellulare. Un paio di anni fa, per descrivere questo processo dinamico, interattivo, coniai il termine topobiologia (dalla parola greca *topos*, che significa luogo).

Come ha chiarito la scoperta delle CAM, uno dei fattori chiave nel determinare la localizzazione di una cellula embrionale (e, in definitiva, la forma e la configurazione dei tessuti) è la presenza di molecole che favoriscono l'adesione tra cellule. Negli ultimi anni è stato dimostrato che diverse famiglie di molecole mediano questa adesione tra cellule e le interazioni intercellulari correlate (si veda in proposito l'articolo *Le molecole che fanno aderire le cellule* di Gerald M. Edelman in «Le Scienze» n. 190, giugno 1984). A parte le CAM, le molecole più importanti che svolgono questa funzione sono le SAM, responsabili dell'adesione al substrato, e le CJM, responsabili delle varie forme di giunzione tra cellule. Mentre le CAM si trovano sempre alla superficie delle cellule e mediano l'interazione tra cellula e cellula, le SAM vengono riversate dalle cellule nell'ambiente extracellulare e formano parte di quella complessa matrice extracellulare a cui spesso si attaccano le cellule. Le CJM contribuiscono a formare strutture complesse tra cui le giunzioni strette o occludenti, le giunzioni comunicanti o intervallate e le giunzioni aderenti, che uniscono le cellule nei tessuti.

Tutte le CAM, le SAM e le CJM note sono proteine complesse, ma le strutture di alcune di esse sono ormai discretamente conosciute. L'attenzione viene oggi concentrata sul modo in cui ciascuna di queste molecole agisce nell'embrione e sulla eventualità che esista un'interazione tra le funzioni delle differenti molecole. Si è così potuto stabilire che praticamente tutte le CAM note si legano l'una all'altra con un meccanismo «omofilo»: una CAM su una cellula si lega a una CAM dello stesso tipo su una cellula giustapposta. Tra le CAM esistono, tuttavia, differenze: esse mostrano differenti specificità di legame e una dipendenza variabile dagli ioni, per esempio dallo ione calcio.

Uno dei motivi ricorrenti dello sviluppo embrionale è la transizione dall'epitelio al mesenchima e viceversa. Il mesenchima è un insieme di cellule non legate tra loro, migranti, le quali non risultano solidamente organizzate entro una configurazione geometrica. L'epitelio è invece una lamina di cellule estremamente organizzate dal punto di vista geometrico e spesso congiunte, alla base del tessuto, alla matrice extracellulare. L'unione dell'epitelio alla matrice avviene per mezzo delle SAM, mentre le cellule epiteliali sono spesso legate tra loro da giunzioni composte da varie CJM. Dato che la transizione da una forma di tessuto a un'altra è un meccanismo basilare dello sviluppo, sarebbe utile conoscere quale di queste molecole sia responsabile della conversione iniziale del mesenchima in lamina epiteliale.

Recenti esperimenti eseguiti nel mio laboratorio suggeriscono che il ruolo iniziale nel legare le cellule epiteliali sia sostenuto dalle CAM e che la formazione di giunzioni come le giunzioni intervallate e le giunzioni aderenti dipenda dal legame delle CAM. Questo risultato è stato ottenuto inserendo, mediante transfezione, DNA responsabile della sintesi delle CAM in cellule coltivate *in vitro*, normalmente prive di questo DNA. Le CAM prescelte erano di due tipi: quella del sistema nervoso (N-CAM) e quella epatica (L-CAM, dall'inglese *liver*, fegato). Esse furono le prime CAM scoperte e vennero denominate con riferimento ai tessuti in cui furono trovate per la prima volta. Oggi si sa che hanno, in realtà, una distribuzione molto più ampia nell'embrione.

Prima della transfezione del gene della L-CAM, le cellule erano separate l'una dall'altra e assomigliavano a un mesenchima. Dopo il trattamento, esse ap-

Distribuzione di cinque CAM (molecole per l'adesione tra le cellule) in un embrione di pollo poco prima della schiusa. Ogni CAM ha una diversa specificità di legame e la loro distribuzione fa pensare che siano tutte coinvolte nella morfogenesi. La disposizione delle CAM varia nel tempo: in una fase più precoce le aree si sovrapponevano coprendo la maggior parte della superficie disegnata.

parivano unite a formare una lamina che aveva alcune delle caratteristiche degli epiteli. Nella lamina comparivano tra le cellule giunzioni sia intervallate sia aderenti. Sembra probabile che l'azione dei geni delle CAM faccia riunire le CJM già esistenti in giunzioni: questo processo sarebbe, di conseguenza, CAM-dipendente. Inoltre abbiamo osservato che l'aggiunta di frammenti di anticorpi contro la L-CAM, bloccando il tipico legame omofilo di queste molecole, produce la disintegrazione delle lamine e una netta diminuzione del numero di giunzioni tra le cellule.

Questi risultati lasciano intravvedere quale sia il significato delle CAM nello sviluppo dell'embrione. Come ho già fatto notare prima, la transizione dall'epitelio al mesenchima e viceversa è uno degli eventi fondamentali dello sviluppo embrionale; tuttavia non è l'unico. Tra gli altri si annoverano certi movimenti precisi di tessuti e la formazione di nuovi limiti tra questi ultimi, processi che fanno tutti parte del quadro complessivo della morfogenesi, in cui compaiono nell'animale organi e tessuti distinti e specializzati. Il ruolo significativo delle CAM in questi eventi spaziali è dimostrato dal fatto che CAM con differenti specificità sono distribuite diversamente e ciascuna in un modo caratteristico che si afferma nel tempo e nello spazio nell'embrione in via di sviluppo.

Un ottimo esempio di come si svolgano tutti questi processi topobiologici è offerto dallo sviluppo delle penne, una fase fondamentale della crescita dell'embrione di pollo. Le penne si formano tramite una serie di processi coordinati, che includono il movimento, la divisione e la morte di cellule secondo schemi prefissati, in cui le cellule aderiscono l'una all'altra e si differenziano. In uno stadio precoce, il coordinamento dello sviluppo dipende in buona parte dalle interazioni tra due tipi di tessuti embrionali, mesoderma ed ectoderma. Dal primo deriva la maggior parte delle ossa e dei muscoli, dal secondo la cute e il sistema nervoso.

In alcune parti dell'embrione di pollo in via di sviluppo, uno strato di tessuto ectodermico specializzato, l'epidermide, poggia sopra tessuti di origine mesodermica. Il primo passo nella formazione delle penne è l'induzione nell'epidermide di strutture particolari, le papille ectomesodermiche, che successivamente daranno origine alle penne. La comparsa di tali strutture implica il passaggio di un insieme complicato di segnali chimici tra mesoderma ed ectoderma, un processo conosciuto come induzione embrionale. L'efficacia di questi segnali dipende a sua volta dall'attività delle CAM, che può essere studiata trasferendo tessuti cutanei embrionali in colture *in vitro*.

Le CAM tengono unite le cellule l'una all'altra tramite un meccanismo omofilo: una CAM su una cellula si lega allo stesso tipo di CAM presente su una cellula giustapposta (*schema in basso*). Le immagini in alto sono state ottenute ombreggiando le CAM con platino e osservandole al microscopio elettronico. L'una (*a sinistra*) mostra tre molecole che favoriscono l'adesione delle cellule nervose (N-CAM), unite da un «pezzo di raccordo» costituito dalle aree che dovrebbero formare una protuberanza attraverso la membrana cellulare esterna. L'altra (*a destra*) mostra una singola molecola che favorisce l'adesione delle cellule epatiche (L-CAM). Le CAM appaiono come molecole lineari munite di cardine. Lo schema in basso illustra due ipotetici meccanismi di legame basati sulla struttura della N-CAM.

Recenti esperimenti effettuati nel mio laboratorio hanno fatto luce sulle relazioni esistenti tra il legame delle CAM e la trasmissione di segnali che conduce a una distribuzione ordinata delle papille ectomesodermiche. In parole povere, la tattica degli esperimenti è consistita nell'intralciare il legame delle CAM e nell'osservarne l'effetto sulla complessa geometria normale delle papille ectomesodermiche. Nel corso di uno sviluppo normale, le papille indotte cominciano ad apparire in corrispondenza della linea mediana dell'embrione e formano una disposizione esagonale a mano a mano che si diffondono lateralmente. Questa distribuzione compare sia nei placodi, precursori delle papille ectomesodermiche che formeranno le strutture della penna matura, sia negli addensamenti dermici, cioè nelle cellule mesodermiche sottostanti, che forniscono il segnale di induzione necessario alla formazione dei placodi nell'epidermide sovrastante.

Il fondamento logico dell'esperimento era che le cellule dei placodi epidermici fossero legate dalla L-CAM, ma non presentassero N-CAM. Per contro, gli addensamenti dermici erano tenuti uniti dalla N-CAM e non contenevano affatto L-CAM. Affinché potesse stabilirsi una configurazione appropriata, abbiamo suggerito la necessità di un circuito di segnalazione prodotto dalle CAM. In questo circuito i segnali sarebbero trasmessi non soltanto dal mesoderma all'epidermide, ma anche dall'epidermide al mesoderma. Pertanto abbiamo introdotto nella coltura anticorpi diretti contro le L-CAM, i quali potevano influire soltanto sui legami tra cellule epidermiche e non su quelli degli addensamenti dermici. Malgrado ciò, si è osservato che, a seguito dell'introduzione dei suddetti anticorpi, avveniva un mutamento anche nella configurazione degli addensamenti dermici (*si veda l'illustrazione in alto a pagina 39*).

Questi risultati hanno dimostrato che, modificando i legami delle CAM in uno strato cellulare, si può alterare il circuito di segnalazione da cui dipende l'intera configurazione. Sia che si alteri la risposta delle cellule epidermiche ai segnali che provengono dagli strati sottostanti, sia che vengano modificati i messaggi diretti verso il basso il risultato è lo stesso: non si forma una configurazione normale. Inoltre, questo effetto è duraturo. Anche dopo che gli anticorpi sono stati eliminati e le cellule modificate sono state tenute in coltura per 10 giorni, il normale sviluppo delle penne ha continuato a essere gravemente perturbato. Come indicano questi esperimenti, la comparsa di una configurazione corretta dipende dalla correlazione tra legami delle CAM e le risposte cellulari ai segnali di induzione.

Questa correlazione mantiene il suo significato a mano a mano che procede lo sviluppo delle penne. La penna matura è una struttura complessa, con un rachide, o asse di sostegno, centrale da cui partono le barbe, assi di sostegno più sottili tenuti insieme da filamenti ancora più sottili, le barbule. Tutte queste strutture sono costituite da una proteina fibrosa, la cheratina, che si conserva quando, alla fine del processo di maturazione, le cellule che formano la penna muoiono. Barbe e barbule si originano nel germe della penna da ispessimenti epidermici (*si veda l'illustrazione nella pagina successiva*). Tali ispessimenti, dopo essere comparsi come rilievi della superficie interna del germe della penna, si allungano e danno origine a strutture colonnari che si interconnettono: queste, cheratinizzandosi e separandosi in filamenti, formano un ciuffo di barbe le quali alla fine del processo risultano tenute insieme dalle barbule.

Le cellule che costituiscono gli ispessimenti epidermici si formano da precursori che, in origine, esprimono sia la L-CAM sia la N-CAM. Tuttavia, a mano a mano che gli ispessimenti si sviluppano, compare un'interessante specializzazione, che distingue la parte rilevata da quella infossata. Con il procedere della maturazione della penna, le cellule degli ispessimenti esprimono soltanto la L-CAM; a un certo punto, le cellule degli avvallamenti cessano di produrre queste proteine, esprimono soltanto la N-CAM e proliferano, il che provoca la formazione di zone, le piastre marginali, che si affiancano a ogni ispessimento epidermico. Poco dopo, tutte le cellule che esprimono la L-CAM cominciano a sintetizzare cheratina, mentre quelle che esprimono la N-CAM si estinguono. Di conseguenza le linee di confine tra zone di espressione della N-CAM e zone di espressione della L-CAM si trasformano nelle strutture colonnari che daranno origine a barbe e barbule.

In questo modo, l'espressione di CAM specifiche, associata al differenziamento e alla morte di determinate cellule, può condurre alla morfogenesi. Tuttavia, sarebbe erroneo pensare che le CAM siano le uniche molecole morforegolatrici. Un esempio, che si basa su recenti esperimenti eseguiti nel mio laboratorio, mostra come anche le SAM possano influire sugli eventi di modellamento dell'embrione. Esso riguarda un gruppo di cellule che danno origine ai gangli della radice dorsale (la radice dorsale è un fascio di nervi che penetra in ogni vertebra nell'organismo maturo). Le vertebre stesse si sviluppano da tessuto mesodermico che si raggruppa in segmenti, i somiti, ciascuno dei quali corrisponde a una singola vertebra. I somiti danno origine anche a un buon numero di altri tessuti: per esempio, gli addensamenti mesodermici della penna provengono da mesenchima che ha origine in essi. I gangli della radice dorsale si formano da cellule ectodermiche che migrano come mesenchima da una struttura particolare, la cresta neurale.

Perché possa formarsi la corretta configurazione gangliare, le cellule della cresta neurale devono penetrare soltanto nella porzione mesenchimale del somite (lo sclerotomo), che si trova nella parte frontale. In che modo viene guidato questo ingresso? Studi preliminari, condotti da Jean-Paul Thiery del Conseil National de la Recherche Scientifique (CNRS) a Parigi, avevano dimostrato che le cellule della cresta neurale migrano lungo percorsi segnati da varie molecole, tra cui la fibronectina; e la fibronectina è una SAM. Inoltre, i miei colleghi e io avevamo scoperto che un'altra SAM, la citotactina, si distribuisce seguendo schemi precisi durante l'embriogenesi. La citotactina è in grado di legarsi alla fibronectina come pure a una terza SAM, il proteoglucano che si lega alla citotactina (CTBP, dall'inglese *cytotactin-binding proteoglycan*). A questo punto è sembrato opportuno stabilire se l'ordinamento periodico impartito dalle cellule della cresta neurale quando migrano nello sclerotomo fosse correlato

Le CAM modificano la forma di gruppi di cellule, come si può osservare in queste microfotografie scattate nel laboratorio dell'autore. Le cellule in coltura che sono prive del gene per la L-CAM assumono un aspetto incoerente simile a quello che si nota nel mesenchima dell'embrione (*a sinistra*). Quando il gene per la L-CAM viene introdotto in esse e attivato per produrre questa molecola, le cellule assumono una disposizione più regolare, che ricorda quella di un epitelio (*al centro*). Tecniche basate sulla fluorescenza mettono in evidenza la localizzazione della L-CAM sulla superficie della cellula (*zone chiare nella microfotografia a destra*).

con la distribuzione delle tre SAM nei somiti.

Un indizio interessante è stato offerto dall'esame della distribuzione delle SAM durante l'invasione dei somiti da parte delle cellule della cresta neurale. In stadi precoci dell'embriogenesi, la citotactina, la fibronectina e il CTBP erano distribuiti uniformemente sull'intera estensione del somite, ma con il procedere dello sviluppo è comparso un ordinamento periodico: benché la fibronectina fosse ancora più o meno uniformemente diffusa in tutto il somite, la citotactina era già concentrata nella parte anteriore e, successivamente, il CTBP appariva addensato nella parte posteriore. Questa distribuzione alterna delle SAM, che compariva non appena le cellule migranti penetravano nello sclerotomo, era provocata da molecole sintetizzate dalle cellule del somite e non dalle cellule in arrivo dalla cresta neurale.

Esperimenti separati sono stati effettuati nel tentativo di distinguere gli effetti delle SAM sulla forma e sul movimento delle cellule. Ricerche eseguite su colture *in vitro* hanno dimostrato che la citotactina e il CTBP fanno assumere alle cellule nervose una forma un poco più arrotondata, che è caratteristica di cellule che non migrano. Inoltre, queste cellule arrotondate non migrerebbero in regioni dove sono presenti le SAM. Coltivate su un intreccio di sole molecole di fibronectina, le cellule della cresta neurale sono apparse, invece, più piatte e hanno migrato con facilità, mentre hanno mostrato un comportamento intermedio su miscugli di fibronectina e dell'una o dell'altra delle due SAM. Benché rimangano da chiarire molti particolari, la conclusione generale che si può trarre da questo lavoro è che vari miscugli delle differenti SAM legate l'una all'altra formano intrecci che possono avere differenti effetti sul comportamento e sulla migrazione di cellule all'interno dell'embrione.

È sorprendente che questi avvenimenti mediati dalle SAM risultino coordinati topologicamente con l'espressione delle CAM: le cellule della cresta neurale, quando migrano, perdono la N-CAM in superficie, ma quando interagiscono nel somite per formare i gangli della radice dorsale presentano di nuovo queste proteine sulla loro superficie. (Infatti, le molecole di N-CAM sono presenti sulla maggior parte delle cellule nervose allo stadio completamente maturo, differenziato.) Chiaramente, durante lo sviluppo, CAM e SAM sono regolate in modo coordinato; entrambi i tipi di molecole possono influire sulla configurazione dei tessuti. Ho già detto

La penna si forma da un precursore chiamato placode. Cellule mesenchimali mesodermiche migrano al di sotto di una semplice lamina di cellule epidermiche (*a*) e formano un insieme che invia segnali verso l'alto per indurre la formazione del placode. La lamina epidermica si rigonfia formando una papilla (*b*) e successivamente un cilindro di epidermide (*c*). All'interno del cilindro si formano ispessimenti colonnari (*d*) che, più tardi, si suddividono per formare la struttura ramificata caratteristica della penna matura (*e-l*).

che, prima che le CJM possano formare giunzioni tra loro, probabilmente le cellule devono essere legate dalle CAM; sembra quindi che in generale, quando un organismo si sviluppa, entrino in gioco subordinazioni complesse tra queste tre classi di molecole morforegolatrici.

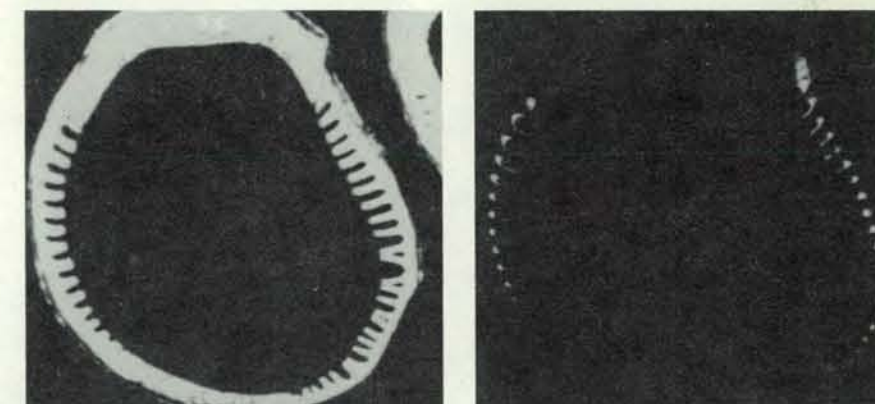
Eppure queste interazioni, per quanto siano complesse, non sono sufficienti a spiegare come si sviluppi l'embrione. Il differenziamento dei tessuti embrionali in tessuti maturi richiede che vengano prodotte, da parte dei tessuti embrionali, proteine specifiche: proteine del fegato nelle cellule epatiche, proteine dei muscoli nelle cellule muscolari e così via. I geni che codificano per queste proteine sono separati dai geni che codificano per le molecole che fanno aderire le cellule. Ma le attività di queste due classi di geni non sono indipendenti: per spiegare il ciclo che le collega ho proposto quindi la cosiddetta ipotesi dei morforegolatori.

Nell'ipotesi dei morforegolatori le cellule sono sottoposte a un controllo chimico e meccanico da cicli di espressione delle CAM e da reti in cui interagiscono le differenti SAM. Un buon esempio di questo controllo è offerto dai cambiamenti di forma e di movimento delle cellule della cresta neurale quando si legano a particolari SAM. L'espressione dei geni per le CAM e per le SAM altera la forma di tali cellule e la loro risposta a determinate configurazioni di segnali; il processo avviene tramite il controllo della formazione di comunità cellulari che, in una data localizzazione, si scambiano segnali. Con questi mezzi, l'espressione dei geni per le CAM e le SAM influisce sull'espressione di altri geni, tra cui i geni di istoregolazione, che codificano per proteine specifiche dei tessuti. In topobiologia, l'interazione tra superfici cellulari controlla le forze chimiche e meccaniche creando comunità cellulari, i cui segnali vengono quindi alterati insieme al loro stato di differenziamento. Questa regolazione dello stato cellulare da parte delle CAM e delle SAM deve svolgere un ruolo fondamentale nell'evoluzione della forma animale e nello sviluppo della configurazione dei tessuti.

L'ipotesi dei morforegolatori ha fatto convergere l'attenzione dei miei colleghi e mia sui geni che codificano per le proteine morforegolatrici, in particolare per le CAM. Bruce A. Cunningham e io abbiamo definito la struttura dei geni per la N-CAM e la L-CAM, determinando le sequenze di DNA responsabili. La nostra ricerca ha dimostrato che entrambe le proteine sono specificate ciascuna da un singolo gene e che i due geni sono sostanzialmente diversi nei particolari della loro struttura, il che fa pensare che le due molecole non siano poi molto affini sotto il profilo evolutivo.

Le differenze tra gene per la L-CAM e gene per la N-CAM stanno in parte nel modo in cui i geni vengono trasformati per trasmettere il messaggio definitivo che specifica una proteina. Quasi tutti i geni dei mammiferi contengono regioni codificanti, gli esoni, e regioni non codificanti, gli introni. Dopo che l'RNA è stato trascritto dal gene, le porzioni non codificanti sono rimosse e le altre vengono saldate assieme a costituire l'RNA messaggero (m-RNA) maturo. In alcune circostanze, diverse combinazioni di saldatura dell'm-RNA possono dare differenti proteine. È questo il caso della N-CAM: i suoi 19 o più esoni possono essere saldati assieme in svariati modi, alcuni dei quali danno origine a CAM che differiscono leggermente nella regione molecolare che serve da attacco alla membrana cellulare. Se questa regione subisce cambiamenti può essere alterata la forza del legame che tiene unite le molecole della N-CAM l'una all'altra e al citoscheletro, cioè allo scheletro interno della cellula. D'altro canto, sembra che il gene per la L-CAM, pur comprendendo numerosi esoni, non possa in alcun modo subire saldature alternative.

Si dovrebbe mettere in rilievo il fatto che le forme alternative di saldatura degli esoni, viste nel caso del gene per la N-CAM, non alterano la specificità del legame omofilo delle molecole di N-CAM: queste si uniscono l'una all'altra

L'esperimento di perturbazione mostra che, nella formazione della penna, sono necessari legami tra le CAM. Quando si sviluppa il placode, le cellule mesodermiche sottostanti assumono una disposizione simmetrica a bolle (*1*) e, in seguito, subiscono altri cambiamenti (*3*). Se la coltura *in vitro* avviene in presenza di anticorpi contro la L-CAM, la disposizione a bolle viene alterata; prima si ottengono delle strisce (*2*) poi la forma globale è sconvolta (*4*). Gli anticorpi influiscono solo sui legami della L-CAM (tra cellule epidermiche), e non su quelli della N-CAM (tra cellule mesodermiche). Malgrado ciò la configurazione mesodermica viene distrutta, il che implica l'esistenza di un sistema di segnali tra questi tessuti.

Via via che gli ispessimenti epidermici che daranno origine alle barbe si sviluppano, si nota l'alternanza di una configurazione tipica della L-CAM e di una tipica della N-CAM. Ogni microfotografia mostra una sezione trasversale del follicolo di una penna in via di sviluppo, trattato con anticorpi fluorescenti diretti contro una CAM specifica. La L-CAM (*a sinistra*) unisce le cellule delle colonne, mentre la N-CAM (*a destra*) unisce le cellule presenti negli avvallamenti tra una colonna e l'altra. Poco dopo questo stadio, le cellule con la L-CAM producono una proteina fibrosa, la cheratina, mentre le cellule con la N-CAM scompaiono.

indipendentemente dal modo in cui sono stati saldati i loro geni. Solo la forza globale del legame fra le molecole viene influenzata dalle saldature alternative a cui può andare soggetto il messaggio genetico. Ciò avviene perché il numero e la disposizione delle CAM, e quindi l'efficacia complessiva del loro legame, si modificano alterando il modo in cui queste molecole si attaccano alla membrana cellulare. Questa situazione concorda molto bene con l'idea che vi sono forse solo poche decine di CAM con differenti specificità di legame la cui regolazione dinamica in corrispondenza della superficie cellulare da parte di svariati fattori può dare luogo a una gamma molto ampia di sfumature di legame.

Altri membri di questa piccola famiglia di molecole hanno già fatto la loro comparsa in laboratorio. Non molto tempo dopo che Cunningham e io avevamo riferito i nostri risultati, diverse ricerche hanno indicato che la N-CAM è simile a un'altra proteina che si trova nei tessuti del sistema nervoso. Attualmente si ha l'impressione che questa molecola, una glicoproteina associata alla mielina (MAG, dall'inglese *myelin-associated glycoprotein*), sia anch'essa una CAM. Successive ricerche sul DNA, effettuate da Masotoshi Takeichi e collaboratori dell'Università di Kyoto, hanno dimostrato che esistono diverse altre CAM, le caderine, analoghe alla L-CAM per la sequenza del DNA e la struttura, ma aventi una distribuzione diversa nei tessuti embrionali.

Tutti questi dati sulla struttura genetica delle CAM erano molto interessanti, ma forse il risultato più sorprendente è stata la scoperta che la N-CAM ha alcune omologie strutturali con le molecole degli anticorpi. Quando, nel 1969, i miei colleghi e io abbiamo completato per la prima volta l'analisi della sequenza amminoacidica dell'intera molecola di un anticorpo, sono risultate evidenti numerose interessantissime relazioni strutturali, genetiche ed evolutive (si veda l'articolo *Struttura e funzione degli anticorpi* di Gerald M. Edelman in «Le Scienze» n. 27, novembre 1970). Queste relazioni erano incluse nella «ipotesi dei dominî», secondo la quale le immunoglobuline constano di due tipi di subunità strutturali e funzionali, ciascuna con una lunghezza corrispondente a circa 100 amminoacidi. Sono questi i dominî variabili, o V, che, in una data classe funzionale, differiscono da molecola a molecola, al contrario di quanto avviene per i dominî costanti, o C.

La caratteristica forma a T della molecola di anticorpo deriva da un assemblaggio specifico di dominî V e C. Ogni anticorpo include una coppia di catene leggere, che sono composte da un dominio variabile e uno costante, e una coppia di catene pesanti, che sono costituite da un dominio variabile e fino a tre dominî costanti. All'interno di questo insieme di catene, i dominî V sono responsabili del legame con l'antigene, mentre i dominî C svolgono un buon numero di funzioni «effettrici», come quella di promuovere l'ingestione di una proteina antigenica estranea legata da parte di specifiche cellule del sistema immunitario, i macrofagi. Fatto assai interessante, tutti i dominî, sia C sia V, delle molecole di anticorpo presentano alcune omologie nella sequenza amminoacidica ed è stata avanzata l'ipotesi che essi si siano formati nel corso dell'evoluzione come risultato di una duplicazione ripetuta di un gene precursore, che specificava una regione non più lunga di un dominio.

Dal 1969, quando venne formulata, l'ipotesi dei dominî è stata ampiamente confermata. In effetti, la famiglia evolutiva così definita è stata ampliata per includervi altre molecole; tra queste vi sono i recettori del fattore di crescita e gli antigeni di istocompatibilità, che mediano alcune interazioni d'importanza fondamentale tra le cellule del sistema immunitario. Questo «clan» evolutivo è stato definito superfamiglia delle immunoglobuline. A mano a mano che venivano scoperti nuovi membri della superfamiglia si presentava lo stesso interrogativo: come ha potuto comparire nel corso dell'evoluzione questo gruppo così diversificato?

Il fatto che la N-CAM fosse omologa alle immunoglobuline ha suggerito una possibile risposta a questo problema. Ho proposto che l'intero sistema immunitario adattativo, caratterizzato dalla presenza della superfamiglia delle immunoglobuline, derivi da un sistema più antico sviluppatosi per favorire l'adesione tra cellule. Il ragionamento su cui poggia quest'ipotesi si basa a sua volta sul fatto che il sistema immunitario adattativo si trova soltanto nei vertebrati, il che sottintende che sia comparso piuttosto tardi nel corso dell'evoluzione, mentre il sistema delle CAM sembra essere molto più diffuso nel regno animale.

Alcune osservazioni recenti offrono un valido sostegno a questa ipotesi. Thomas C. Kaufman e Mark Seeger dell'Indiana University hanno trovato, nel complesso genico *antennapedia* del moscerino della frutta (*Drosophila melanogaster*), una sequenza di DNA omologa a circa due dominî e mezzo della sequenza genica della N-CAM; questa sequenza codifica per una proteina che è stata denominata amalgama. Successivamente, Corey S. Goodman e collaboratori dell'Università della California a Berkeley hanno mostrato che la sequenza di una ipotetica CAM, chiamata fasciclina II, che è stata trovata nelle cellule nervose di *Drosophila*, è anch'essa omologa alla N-CAM. I moscerini della frutta non hanno nulla che possa far pensare a un sistema immunitario adattativo. Poiché gli insetti hanno molecole simili alla N-CAM e solo i vertebrati possiedono un sistema immunitario basato sugli anticorpi, sembra probabile che i geni delle CAM, presenti in qualche antico precursore sia degli insetti sia dei vertebrati, abbiano offerto la base per la comparsa delle molecole del sistema immunitario adattativo.

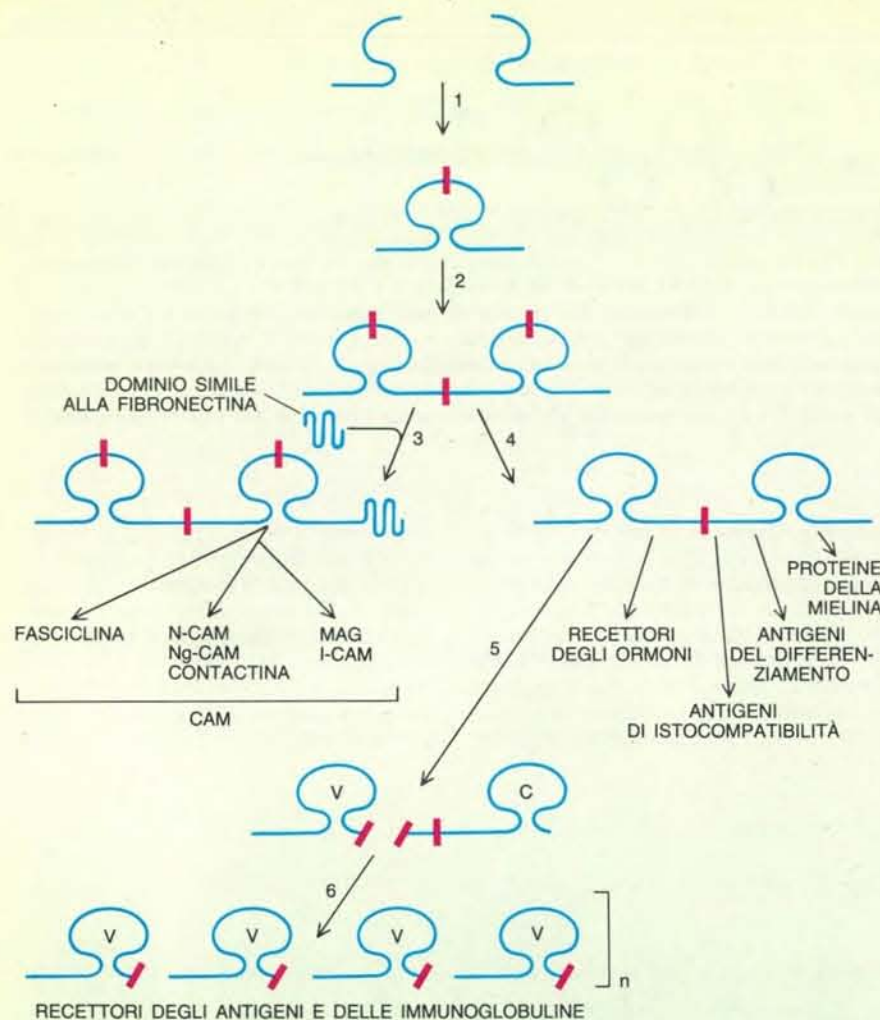
Vari altri fatti contribuiscono a fornire uno schema, che rappresenta un'estensione soddisfacente dell'ipotesi dei dominî, per l'evoluzione delle molecole che favoriscono l'adesione tra cellule e delle loro discendenti. Diversamente dalle CAM, ogni immunoglobulina viene specificata da geni multipli per le regioni V e C, comparsi in seguito a duplicazione di un precursore originale nel corso dell'evoluzione; queste famiglie di geni duplicati mostrano omologie tra le varie specie di vertebrati. Sembrerebbe, però, che queste omologie tra specie possano venire distrutte da mutazioni indipendenti che si accumulano in differenti geni di una simile famiglia di geni duplicati, a meno che non entrino in funzione speciali meccanismi. Quali potrebbero essere questi meccanismi?

Nel 1969 il mio collega Joseph A. Gally e io abbiamo avanzato l'ipotesi che le famiglie a cui appartenevano tali geni potessero evolversi in parallelo nelle diverse specie, se i geni in questione subivano un processo di cosiddetta «conversione genica democratica». I particolari di una simile conversione esulano un poco dalle finalità di questo articolo, ma l'essenza è che i geni di una famiglia si ricombinano l'uno con l'altro in un modo particolare, che rende possibile la coevoluzione. Un fatto ancora più importante è che le famiglie di geni possono servire da «rete di mutazione» che consente la diffusione di mutazioni favorevoli tra i vari membri della famiglia. Poiché una mutazione favorevole per il sistema immunitario adattativo di una data specie può benissimo esserlo anche per il sistema di un'altra specie, queste

La struttura della N-CAM è stata dedotta grazie alle conoscenze sulla sua biochimica e sulla sequenza del DNA nel gene che la specifica. La N-CAM esiste in diverse varianti, le quali differiscono nel numero di esoni (regioni codificanti) del gene per la N-CAM che esse incorporano. Si vedono qui 19 esoni; altri ne sono stati scoperti di recente. Tutte le varianti note includono cinque anelli chiusi da legami disolfuro (*a sinistra*). Lo schema mostra tre varianti: *ssd* si ferma all'esone 15 e non possiede una regione transmembrana; *sd* include gli esoni 17 e 19, ma manca del 15; infine *ld* include l'esone 18, ma è priva dell'esone 15.

STORIA DEI LEGAMI CONTROLLATI DALLE CAM IN SEGUITO A MOVIMENTO DI CELLULE E LAMINE E AD ACCRESCIMENTO

L'ipotesi dei morforegolatori cerca di spiegare in che modo le CAM e le SAM (molecole per l'adesione al substrato) producano comunità cellulari, interagendo poi con esse nel determinare la forma del corpo di un animale. Nei cicli controllati dalle CAM (*a sinistra*), le comunità cellulari tenute insieme da queste molecole producono i morfogeni, che inducono le cellule legate da CAM differenti ad alterare l'attività dei loro geni morforegolatori (*MR*), geni che codificano per le CAM e le SAM. I morfogeni influiscono anche sui geni istoregolatori (*HR*), geni che codificano per proteine specifiche dei tessuti. In alcune cellule, i geni *HR* sono regolati a loro volta da geni selettori (*S*). Le comunità cellulari, interagendo, alterano il loro ambiente: entrano allora in gioco nuovi cicli delle CAM (*sopra a destra*), che si inviano reciprocamente segnali (*doppie frecce*).

La superfamiglia delle immunoglobuline che, come è ora noto, include le CAM consta di molecole che sono strutturalmente e funzionalmente diverse. Ciascuna include diversi domini. L'IgG, la molecola che costituisce gli anticorpi, circola nei liquidi organici e si lega a sostanze estranee. Le proteine di classe I del maggior complesso di istocompatibilità (MHC), componenti di quasi tutte le cellule animali, si associano con antigeni trasformati e servono per presentare in modo efficace gli antigeni ai linfociti *T*. La corta catena delle proteine di classe I dell'MHC ha un unico dominio, la microglobulina beta$_2$. La MAG è una molecola che si ritiene abbia un ruolo nell'adesione tra cellule e che si trova nelle membrane delle cellule gliali, cellule non eccitabili che fanno parte del sistema nervoso. Le cellule gliali (e anche la MAG) prendono parte alla mielinizzazione di determinati neuroni.

Lo schema evolutivo mostra come si è formato il sistema immunitario adattativo a partire da antiche CAM. Gli esoni codificanti per due «mezzi domini» si sono combinati nel formare il gene per un singolo dominio, un'unità funzionale di una proteina (*1*). La duplicazione genica ha prodotto il gene per una proteina con molti domini, una antica N-CAM (*2*). Un «rimaneggiamento di esoni» ha congiunto un dominio simile alla fibronectina a quelli per la famiglia della N-CAM (*3*). La duplicazione genica e la divergenza hanno dato origine a una varietà di CAM. Altri geni hanno perduto gli introni (segmenti di DNA non codificante) all'interno dei domini, ma non quelli tra un dominio e l'altro (*4*). In questo gruppo, una successiva divergenza ha dato origine ad altri membri della superfamiglia delle immunoglobuline. L'inserzione di un elemento genetico (con tutta probabilità proveniente da un virus) ha fornito la possibilità di una diversificazione proteica in differenti linfociti (*5*). Insieme a questo cambiamento si è avuto un aumento nel numero delle regioni variabili (*V*), il che ha costituito la base per la produzione di una vasta popolazione di recettori per gli antigeni e gli anticorpi durante l'esistenza di un individuo (*6*).

reti di mutazione, associate alla pressione selettiva, potrebbero anche contribuire a conservare le omologie nelle famiglie dei geni per le immunoglobuline. Quest'idea, oltre a essere valida per gli anticorpi, ha anche un'importanza generale e si può applicare all'evoluzione di tutte le famiglie multigeniche.

La cooperazione funzionale tra CAM e anticorpi appare oggi evidente nel sistema immunitario, nel quale diversi linfociti, che sono tra le cellule effettrici più importanti di questo sistema, devono aderire al bersaglio per poter effettuare le loro funzioni. Recentemente Timothy A. Springer della Harvard Medical School ha scoperto in una varietà di cellule una molecola che è stata denominata I-CAM. Essa è omologa alla N-CAM e, come ha dimostrato Springer, si lega a una molecola, la LFA-1, che si trova nei linfociti. La LFA-1 rassomiglia alle integrine, molecole della superficie cellulare che hanno la funzione di recettori per le SAM.

Se tutte queste nozioni vengono combinate in uno schema generale, emerge un magnifico esempio di opportunismo evolutivo. Il punto di partenza è un filamento di DNA, corrispondente per dimensione alla metà circa di un dominio della N-CAM. L'analisi della struttura del gene per la N-CAM rivela che questo precursore ridotto si è duplicato, dando origine a geni simili a quello per la N-CAM negli antenati degli insetti e dei vertebrati. Quindi, tramite un rimaneggiamento genetico di esoni, sono state introdotte regioni di DNA che assomigliano a quelle che codificano per SAM come la fibronectina. Il singolo gene per molecole di tipo N-CAM, che favoriscono l'adesione, è stato poi a sua volta duplicato e questo fatto ha dato origine a tutte le altre CAM affini, tra cui quelle presenti nel sistema nervoso centrale come la MAG.

Tali molecole sono state tutte utilizzate topobiologicamente nella morfogenesi. Ma in qualche antico vertebrato (o in un suo immediato precursore), un gene della famiglia è stato dirottato verso una finalità interamente differente. Per duplicazione di un segmento di DNA, simile alla sequenza responsabile della sintesi di una CAM, sono comparse le regioni variabili e costanti dei geni delle immunoglobuline, i recettori presenti su certi linfociti e gli antigeni di istocompatibilità. In seguito, i geni per le regioni variabili si sono duplicati ripetutamente dando origine a famiglie mantenute simili per conversione genica democratica. Analoghi eventi di conversione hanno avuto luogo nella famiglia dei geni per gli antigeni di istocompatibilità. CAM come la I-CAM sono comparse per servire da punti di attacco per la LFA-1, il recettore di superficie dei linfociti evolutosi indipendentemente.

Così, le funzioni chiave del sistema immunitario adattativo hanno avuto origine da un antico sistema che favoriva l'adesione tra cellule, il quale ha operato per via topobiologica nel regolare le migrazioni cellulari e le configurazioni tissutali che sono alla base della morfogenesi. I linfociti, pur non intervenendo nella morfogenesi, si basano su interazioni tra cellule regolate in modo complicato. Questa regolazione dipende a sua volta da meccanismi specifici di legame, che possono risalire alla fonte del sistema immunitario adattativo in seno al sistema responsabile dell'adesione tra cellule. Un tipo specifico di ripiegamento della proteina si nota in corrispondenza di tutti i siti di legame delle immunoglobuline, e probabilmente anche di tutti quelli delle CAM. Esso si chiama struttura a lamina pieghettata beta; due lamine di questo tipo formano, nel dominio dell'immunoglobulina, una particolare struttura a sandwich. La lamina pieghettata beta può fare parte di un antico «linguaggio» di legame e regolazione cellulare, che il sistema immunitario adattativo ha ereditato dai suoi predecessori simili alle CAM.

L'interazione tra anticorpo e antigene, che può aver luogo in provetta in un mezzo liquido senza tener conto della posizione, viene spesso considerata emblematica del funzionamento dell'intero sistema immunitario. Tuttavia, non si deve dimenticare che il sistema immunitario consta sia di tessuti solidi (come i linfonodi) sia di tessuti liquidi (come il sangue). Anche tra i linfociti circolanti, liberi, vi sono popolazioni che «si dirigono» preferenzialmente verso particolari tessuti dell'organismo. Benché in provetta possano aver luogo risposte immunitarie di singole cellule, è comunque probabile che, nella risposta immunitaria globale, la localizzazione sia altrettanto critica che nella morfogenesi. In effetti, l'analisi delle molecole morforegolatrici e delle immunoglobuline mostra che questi due sistemi sono uniti da legami profondi sia dal punto di vista genetico sia da quello evolutivo.

Può non essere inopportuno aggiungere un'ultima parola sui filoni di ricerca. La linea di pensiero che unisce le molecole del sistema immunitario e le CAM non avrebbe potuto prendere corpo in un'area qualunque della specializzazione scientifica. Era necessario rincorrere una vaga premessa seguendo un lungo filone di ricerca (sulle CAM) fino alla soluzione di un problema (le origini delle immunoglobuline), che si era presentato al compimento di un precedente filone di ricerca, apparentemente del tutto privo di connessioni con l'altro. Questa linea di pensiero sarebbe risultata molto impoverita se fosse stata concepita entro i limiti di una disciplina specializzata, come l'immunologia o l'embriologia. Solo considerando la biologia nelle sue più ampie prospettive evoluzionistiche, genetiche e di sviluppo, pur perseguendo una ricerca specializzata, è possibile collegare quelle che a prima vista possono sembrare questioni distinte in un insieme che sia nel contempo organico e soddisfacente sotto il profilo intellettuale.

BIBLIOGRAFIA

EDELMAN G. M. e GALLY J. A., *Arrangement and Evolution of Eukaryotic Genes* in *The Neurosciences* a cura di Francis O. Schmitt, The Rockefeller University Press, 1970.

EDELMAN GERALD M., *CAMs and Igs: Cell Adhesion and the Evolutionary Origins of Immunity* in «Immunological Reviews», n. 100, pp. 11-45, dicembre 1987.

TAN SEONG-SENG, CROSSIN KATHRYN L., HOFFMAN STANLEY ed EDELMAN GERALD M., *Asymmetric Expression in Somites of Cytotactin and Its Proteoglycan Ligand Is Correlated with Neural Crest Cell Distribution* in «Proceedings of the National Academy of Sciences of the United States of America», 84, n. 22, novembre 1987.

EDELMAN GERALD M., *Topobiology: An Introduction to Molecular Embryology*, Basic Books, Inc., 1988.

La dinamica dell'aurora polare

Le interazioni tra campo magnetico terrestre e vento solare producono un gigantesco generatore che alimenta le spettacolari aurore polari e che può servire da modello per la comprensione di altri fenomeni astrofisici

di Syun-Ichi Akasofu

Alcuni dei primi osservatori delle aurore boreali ritenevano che queste spettacolari manifestazioni luminose visibili nel cielo delle regioni polari fossero dovute alla rifrazione della luce solare da parte dell'atmosfera, con un meccanismo analogo a quello con cui si forma l'arcobaleno. Secondo le loro ipotesi, le pieghe scintillanti erano causate dai moti atmosferici. Oggi i fisici che studiano il fenomeno sanno che la luce aurorale viene emessa quando atomi e molecole della ionosfera collidono con elettroni provenienti dal Sole. Il movimento dei «drappeggi» aurorali è provocato non dalla turbolenza atmosferica, ma da variazioni dei campi elettromagnetici che accelerano gli elettroni, proprio come l'illusione del movimento su uno schermo televisivo è data da variazioni del campo magnetico che dirige gli elettroni dal tubo catodico sullo schermo.

Nel caso dell'aurora, che cosa funge da tubo catodico? Da dove viene l'energia? E perché questa energia sembra di tanto in tanto fluttuare, facendo sì che l'intensità dell'aurora cresca e poi diminuisca nel cielo polare? Più di 20 anni fa si è accertato che le manifestazioni aurorali avvengono perché la ionosfera è bombardata da fasci di elettroni generati dalla complessa interazione fra il vento solare e il campo magnetico della Terra. La natura geomagnetica dell'aurora può essere osservata chiaramente dallo spazio: centrato su ciascuno dei poli geomagnetici vi è un grande ovale luminoso (l'aurora boreale in corrispondenza del Polo nord e quella australe in corrispondenza del Polo sud), una caratteristica permanente del nostro pianeta.

I miei colleghi e io abbiamo scoperto una relazione quantitativa fra il vento solare e la generazione dell'energia che provoca sia l'aurora, sia altre perturbazioni del campo geomagnetico. Stiamo anche componendo un quadro sempre più preciso dell'influenza dell'attività del Sole sul vento solare. Questi progressi fanno sorgere l'interessante possibilità di sviluppare un metodo per prevedere, con l'aiuto del calcolatore, l'intensità dell'attività aurorale. Dato che le aurore possono interferire notevolmente con le comunicazioni radio normali e con quelle via satellite, oltre che con le linee elettriche e con alcuni sistemi di difesa, la capacità di prevedere l'intensità aurorale diventa sempre più importante via via che le attività umane si estendono alle regioni polari e allo spazio.

Un generatore nella magnetosfera

La comprensione del meccanismo di formazione dell'aurora è stata possibile grazie a una rivoluzione nelle conoscenze scientifiche sull'ambiente magnetico della Terra. Per molto tempo si è supposto che il campo magnetico terrestre fosse essenzialmente un campo di dipolo, come quello di un magnete a barra, in cui le linee di forza del campo magnetico decorrono, dal polo sud al polo nord, simmetricamente rispetto all'asse geomagnetico. Tuttavia la Terra non è immersa nel vuoto; essa è continuamente investita dal vento solare, un plasma diluito di ioni idrogeno (protoni) e di elettroni emessi dalla corona solare.

Il vento solare confina il campo magnetico terrestre in un volume di forma simile a una cometa, la magnetosfera. Sul lato rivolto al Sole il vento solare comprime la magnetosfera fino a una distanza dal nostro pianeta pari a circa 10 volte il raggio terrestre. Sul lato opposto la magnetosfera è confinata in una struttura a forma di manica a vento, la coda magnetosferica, che si estende per più di 1000 volte il raggio terrestre. La magnetosfera contiene plasmi rarefatti di densità e temperatura diverse, derivanti dal vento solare e dalla ionosfera (si veda l'articolo *La coda magnetica della Terra* di Edward W. Hones, Jr., in «Le Scienze» n. 213, maggio 1986).

All'inizio degli anni sessanta i fisici iniziarono a rendersi conto del fatto che il vento solare può estendere e trasportare il campo magnetico della corona fino alle frange più lontane del sistema solare dando così origine al campo magnetico interplanetario. James W. Dungey dell'Imperial College of Science and Technology di Londra ha proposto che questo campo magnetico potrebbe fondersi con le linee di forza del campo geomagnetico che hanno origine nelle regioni polari della Terra. Questo fenomeno, chiamato riconnessione magnetica, avviene con la massima efficienza quando il campo magnetico del vento solare è orientato verso sud, ossia quando è antiparallelo al campo terrestre.

Si riteneva generalmente che la riconnessione fosse un processo stabile, ma una decina di anni fa Christopher T. Russell dell'Università della California a Los Angeles scoprì che le cose non stanno così. Al contrario, si ha la formazione di «matasse» o «corde» di linee di forza, che infine si staccano dalla magnetosfera e sono trascinate nella coda magnetosferica. Le instabilità si formerebbero anche se il campo magnetico interplanetario fosse costantemente orientato verso sud; ma in realtà l'intensità e la direzione del campo mutano continuamente. Le complesse interazioni che ne risultano sono state chiarite in una simulazione al calcolatore da Lou-Chuang Lee dell'Università dell'Alaska a Fairbanks.

Le particelle del vento solare fluiscono lungo il confine della magnetosfera (la magnetopausa) e attraversano le linee del campo magnetico riconnesse. Gli ioni positivi e gli elettroni, avendo carica elettrica opposta, sono deflessi in direzioni opposte (secondo la familiare regola della mano destra) e generano una corrente elettrica (*si veda l'illustrazione a pagina 49*). Questo processo è identico alla produzione di energia in un generatore magnetoidrodinamico. In effetti l'intera magnetopausa costituisce un gigantesco generatore che trasforma l'energia cinetica delle particelle del vento solare in energia elettrica, con una produzione di più di un milione di megawatt. Questo meccanismo alimentato dall'interazione fra vento solare e magnetosfera è il «generatore aurorale».

Correnti allineate con il campo

Il meccanismo del generatore spinge gli iòni positivi verso il lato dell'alba del piano equatoriale della magnetopausa, formando una sorta di terminale positivo; gli elettroni sono deflessi verso il lato del crepuscolo o terminale negativo. Nei plasmi rarefatti permeati da linee del campo magnetico, come quelli presenti nella magnetosfera, gli elettroni si muovono lungo traiettorie elicoidali che si avvolgono intorno alle linee di forza del campo. Il processo con il quale hanno origine queste cosiddette correnti allineate con il campo è stato ricostruito nei dettagli da Akira Hasegawa degli AT&T Bell Laboratories e da altri. Thomas A. Potemra della Johns Hopkins University, Takesi Iijima dell'Università di Tokyo e altri hanno confermato l'esistenza delle correnti mediante magnetometri a bordo di satelliti.

Le correnti allineate con il campo collegano la magnetopausa alla ionosfera. Il lato dell'alba della magnetopausa è così proiettato sulla parte dell'alba dell'ovale aurorale, mentre quello del crepuscolo è proiettato sulla parte del crepuscolo dell'ovale; la parte dell'alba dell'ovale assume perciò una carica positiva e la parte del crepuscolo si carica negativamente. La caduta netta di potenziale nell'ovale è di circa 100 chilovolt.

Le correnti allineate con il campo sono prodotte da elettroni anziché da protoni perché gli elettroni sono più mobili. Nello scenario descritto le correnti di elettroni fluiscono verso il basso nella parte del crepuscolo dell'ovale e verso l'alto nella parte dell'alba. Dato che le aurore sono prodotte dalle collisioni degli elettroni con la ionosfera, ci si potrebbe chiedere perché si osservano manifestazioni luminose nella parte dell'alba dell'ovale. In realtà l'ovale ionosferico è fortemente conduttore e una corrente elettrica fluisce tra i suoi margini interno ed esterno. La corrente torna poi indietro lungo le linee di forza del campo magnetico, generando una corrente secondaria in direzione opposta rispetto alla primaria. Sono le correnti secondarie di elettroni a produrre i fenomeni aurorali sul lato dell'alba.

Una ondulazione che si propaga verso ovest lungo un'aurora attiva appare nel cielo serale sopra Fairbanks, in Alaska. La luce biancastra è emessa da atomi di ossigeno, mentre molecole di azoto ionizzato emettono la luminosità rosea al margine inferiore del drappeggio. A sinistra si osserva un'immagine in falsi colori dell'ovale aurorale sopra il Polo nord, realizzata dal satellite *Dynamics Explorer* da una distanza di tre raggi terrestri. L'immagine, che registra l'emissione degli atomi di ossigeno alla lunghezza d'onda di 130 nanometri, è stata concessa da Louis Frank dell'Università dello Iowa.

La luminosità aurorale si manifesta quando i fasci di elettroni in arrivo subiscono collisioni di alta energia con la ionosfera, eccitando o ionizzando atomi e dissociando molecole (con la formazione di altri atomi allo stato eccitato). Gli atomi eccitati e quelli ionizzati emettono radiazione in un ampio intervallo spettrale (dall'ultravioletto estremo all'infrarosso) via via che gli atomi eccitati ritornano a livelli di energia più bassi e gli ioni si combinano con elettroni liberi.

L'emissione aurorale più comune è una luce verde-biancastra con una lunghezza d'onda di 557,7 nanometri, che è emessa da atomi di ossigeno. Una bella emissione rosa è prodotta da molecole eccitate di azoto. Diversi atomi e molecole presenti nella ionosfera danno origine a emissioni aurorali alle lunghezze d'onda dell'estremo ultravioletto, dell'ultravioletto e dell'infrarosso, che non possono essere osservate da terra perché vengono assorbite dall'atmosfera. Immagini di emissioni nell'estremo ultravioletto, ottenute dal satellite svedese *Viking*, mostrano che l'aurora è sorprendentemente attiva sul lato illuminato dal Sole, spesso di più che sul lato opposto.

Perché la forma a drappeggio?

Vista da terra, l'aurora ha l'aspetto di un drappeggio luminoso a cui sono sovrapposte strutture a raggi. Il drappeggio inizia a una quota di diverse centinaia di chilometri e termina a circa 100 chilometri di altezza, dove l'atmosfera diventa abbastanza densa da arrestare gran parte degli elettroni in arrivo. La struttura ha uno spessore inferiore al chilometro, ma in senso laterale si estende per migliaia di chilometri. Come si può spiegare il fenomeno?

In effetti, il semplice fatto che le correnti di elettroni allineate con il campo riescano a raggiungere la bassa ionosfera è sorprendente. Come avviene per altre particelle cariche nelle fasce di radiazione di Van Allen, il passo della traiettoria elicoidale di un elettrone tende a diminuire via via che l'elettrone si avvicina alla Terra (dove il campo è più intenso). Anzi, il suo moto diventa puramente circolare a quote molto superiori alla ionosfera, e a quel punto l'elettrone è riflesso verso l'alto. Tuttavia le aurore indicano che gli elettroni sono in grado di penetrare in profondità nella ionosfera.

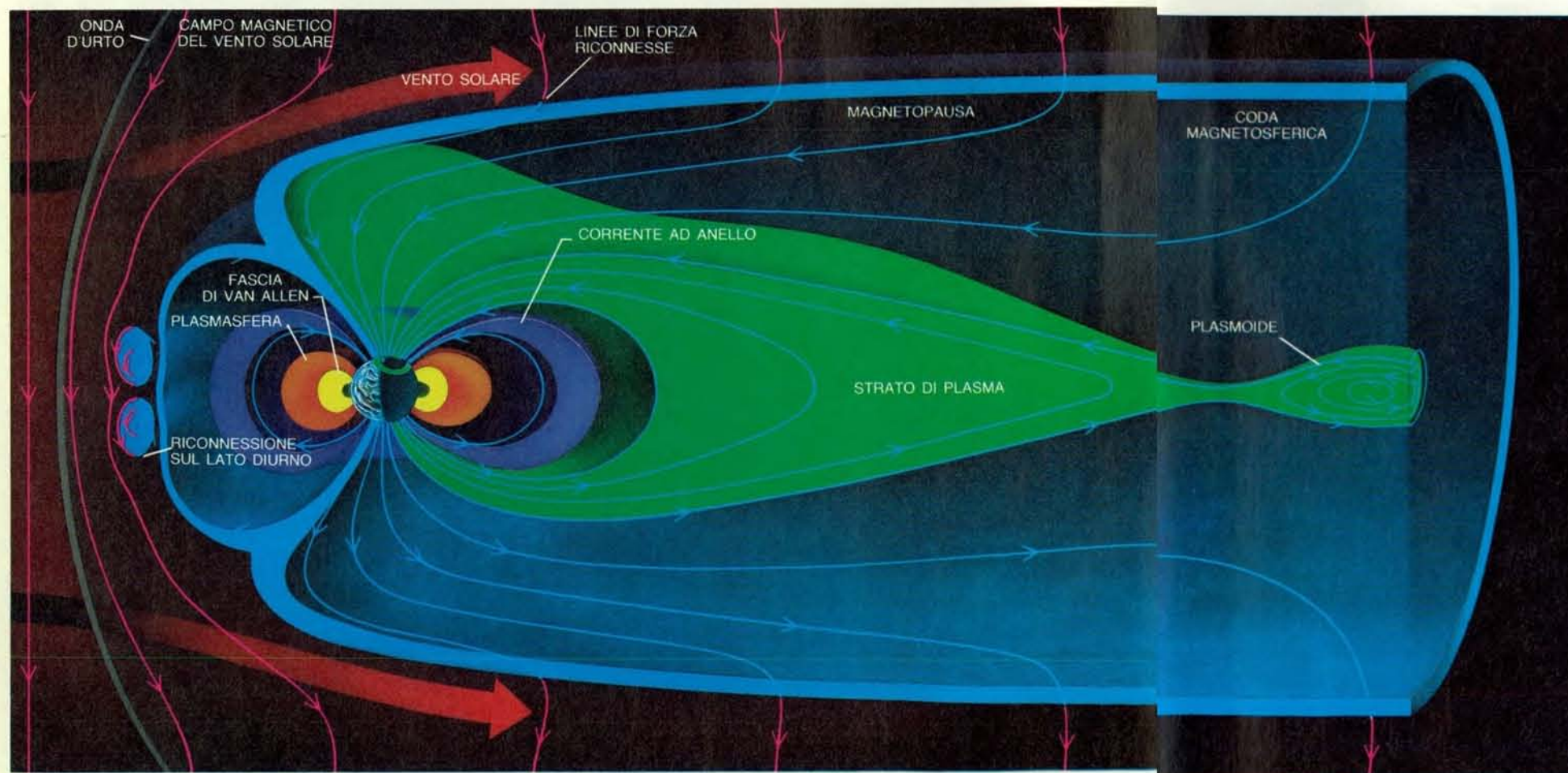
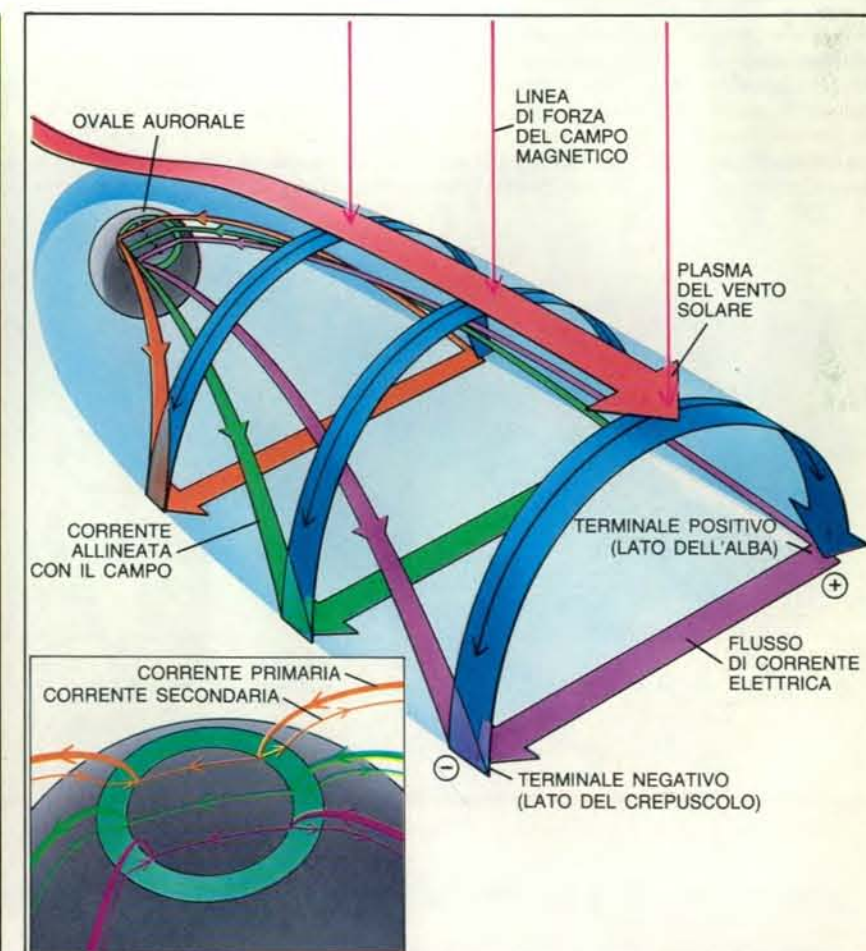
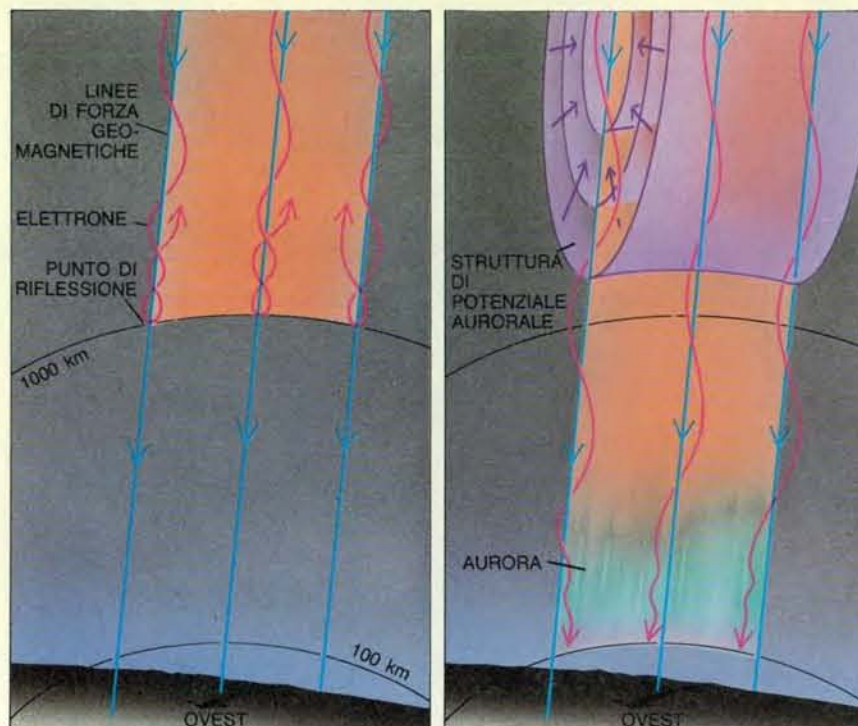
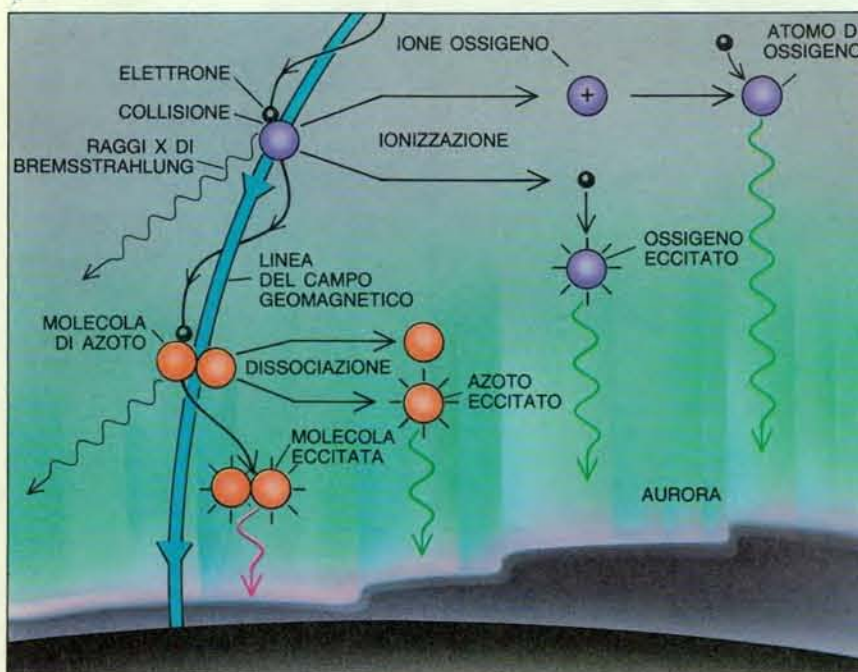
Sembra che il processo cominci quando gli elettroni delle correnti allineate con il campo formano fasci sottili, laminari. Quando nella magnetosfera viene «pompata» abbastanza energia e i fasci laminari raggiungono un'intensità sufficientemente elevata, intorno ai fasci si sviluppa (secondo un meccanismo non ancora chiarito) un particolare campo elettrico, la cosiddetta «struttura di potenziale aurorale», a una quota compresa tra 10 000 e 20 000 chilometri. La regione interna alla struttura sembra suddividersi in strati aventi carica positiva e negativa, che danno origine a un campo elettrico interposto molto intenso (*si veda l'illustrazione a pagina 50, in alto*). Hannes Alfvén dell'Università della California a San Diego è stato il primo a proporre la presenza di un «doppio strato» di questo tipo al di sopra dell'aurora.

Si sa che una struttura simile a un doppio strato si forma alla superficie dell'elettrodo in una lampada a neon, ma l'esatta natura di quella presente nell'alta atmosfera è attualmente oggetto di discussione. Sembra che gli elettroni vengano accelerati verso il basso dal campo elettrico associato al doppio strato. Nel momento in cui arrivano all'estremità inferiore della struttura di potenziale aurorale essi possiedono un'energia di alcune migliaia di elettronvolt, sufficiente per farli penetrare fino a una quota dove l'atmosfera è abbastanza densa da rendere le manifestazioni aurorali visibili a occhio nudo. La conformazione appiattita della struttura di potenziale è correlata alla forma a drappeggio dell'aurora.

I raggi che accompagnano i drappeggi aurorali sono in realtà una serie di vortici che si formano quando campi elettrici antiparalleli (associati alla struttura di potenziale aurorale) lungo la superficie del drappeggio fanno sì che gli elettroni fluiscano in direzioni opposte: si formano mulinelli proprio come accade al confine tra due correnti d'acqua che scorrono in senso opposto. Telecamere ad alta velocità puntate sulla parte inferiore del drappeggio hanno registrato immagini di questi vortici e strutture notevolmente simili sono state generate formando vortici con simulazioni al calcolatore.

L'energia degli elettroni aurorali in relazione alla distribuzione spaziale della struttura di potenziale è stata ampiamente studiata con strumenti a bordo di razzi e satelliti da Louis A. Frank dell'Università dello Iowa, James L. Burch del Southwest Research Institute, Patricia H. Reiff della Rice University e Bengt Hultqvist e colleghi del programma svedese Viking. I ricercatori della Lockheed Missiles and Space Company hanno inoltre confermato che la struttura di potenziale aurorale accelera gli ioni positivi verso l'alto; in effetti, questi ioni rappresentano a volte una parte significativa del plasma magnetosferico.

Nella struttura di potenziale aurorale si hanno anche interazioni tra particelle e onde elettromagnetiche nel plasma, che generano intense onde radio. Donald A. Gurnett dell'Università dello Iowa ha fatto notare che queste emissioni sono così intense che un sistema di rilevamento extraterrestre le individuerebbe molto prima di poter vedere la Terra. Queste emissioni non interferiscono con le trasmissioni radio sulla Terra perché la ionosfera le riflette verso l'alto, proprio come riflette verso il basso le emissioni che hanno origine a terra.

L'esistenza della struttura di potenziale aurorale dovrebbe risultare di grande interesse per gli astrofisici e per i fisici che studiano il Sole, oltre che per quelli che si occupano delle aurore, dato che è ancora opinione comune che un campo elettrico significativo non possa conser-

Il vento solare, un plasma diffuso di protoni e di elettroni emesso dal Sole, confina il campo magnetico terrestre in una struttura a forma di cometa, la magnetosfera. Il vento solare comprime la magnetosfera sul lato rivolto al Sole fino a una distanza di circa 10 raggi terrestri. Sul lato notturno, invece, esso deforma il campo magnetico della Terra, dandogli una forma allungata detta coda magnetosferica che si estende per almeno 1000 raggi terrestri. Il limite della coda magnetosferica è la magnetopausa. Il vento solare possiede un campo magnetico (*in rosso*) che, quando è diretto a sud, come in questo caso, può riconnettersi efficacemente con il campo terrestre (*in blu*). Le particelle del vento solare fluiscono nella magnetosfera lungo le linee di forza riconnesse. Le linee di forza del campo magnetico nel lobo nord della coda puntano verso la Terra, quelle nel lobo sud puntano in direzione opposta. La riconnessione delle linee di forza nella coda magnetosferica può «strizzare» ammassi di plasma, i plasmoidi, che sono espulsi dalla coda magnetosferica.

Il plasma del vento solare fluisce perpendicolarmente alle linee del campo magnetico nella magnetopausa deviando i protoni verso il lato dell'alba della coda magnetosferica e gli elettroni verso quello del crepuscolo. Fra le due regioni fluisce corrente, per lo più direttamente attraverso la coda, ma in parte lungo le linee di forza, da e verso un ampio ovale nella ionosfera. Questa corrente allineata con il campo è costituita da elettroni, che causano l'aurora. Su percorsi paralleli alle correnti primarie si sviluppano correnti secondarie (*riquadro*).

Gli elettroni allineati con il campo seguono traiettorie elicoidali intorno alle linee di forza del campo magnetico. Scendendo verso la ionosfera il passo della traiettoria diminuisce finché essi sono riflessi verso l'alto (*a sinistra*). In certi casi (*a destra*) un campo elettrico particolare, la «struttura di potenziale aurorale», si sviluppa intorno alle correnti di elettroni. Gli elettroni accelerati verso il basso da questa struttura possono scendere in profondità nella ionosfera. L'ampiezza della struttura di potenziale in senso nord-sud è minima, ma essa si estende da est a ovest per migliaia di chilometri e dà all'aurora la forma a drappeggio.

Le molecole e gli atomi nella bassa ionosfera emettono radiazione quando sono colpiti da elettroni accelerati dalla struttura di potenziale aurorale. Gli elettroni rallentati dalle collisioni emettono raggi X di bremsstrahlung (frenamento). Gli urti dissociano le molecole in atomi eccitati che emettono radiazioni nel ricadere a un livello energetico più basso. Altra radiazione è prodotta quando gli elettroni deviati colpiscono ed eccitano gli atomi; questi possono anche venire ionizzati e poi emettere radiazioni nel ricombinarsi con gli elettroni.

varsi lungo le linee di forza del campo magnetico in un plasma rarefatto e che perciò non è possibile accelerare in questo modo particelle cariche. L'osservazione di un campo di questo genere in correlazione con le aurore fa pensare che campi analoghi possano esistere in condizioni astrofisiche come quelle della Nebulosa del Granchio, in cui la densità e l'energia del plasma si avvicinano a quelle della magnetosfera.

Elettrogetti e sottotempeste

L'apporto di energia alla ionosfera da parte del generatore aurorale si compie soprattutto tramite una coppia di correnti elettriche, gli elettrogetti, dirette rispettivamente verso ovest e verso est. Queste correnti si muovono lungo l'ovale aurorale nella bassa ionosfera e generano un intenso riscaldamento che contribuisce alla produzione di venti di grande scala nell'alta atmosfera polare.

Come hanno origine gli elettrogetti? Si ricordi che fra i margini esterno e interno dell'ovale fluiscono correnti elettriche. Queste correnti sono indotte da un campo elettrico che è parallelo alla superficie terrestre e quindi perpendicolare al campo magnetico della Terra ai poli (*si veda l'illustrazione nella pagina a fronte in alto*). In presenza di questi campi tra loro perpendicolari, le particelle cariche vanno incontro a un fenomeno di deriva corrispondente al prodotto vettoriale tra **E** e **B** (dove **E** è il campo elettrico e **B** il campo induzione magnetica), tramite il quale le particelle positive e negative sono deviate nella stessa direzione, dal lato notturno al lato diurno.

Nell'alta ionosfera queste particelle si muovono alla stessa velocità e quindi non vi è corrente netta; però esse impartiscono quantità di moto alle particelle neutre, contribuendo ulteriormente a produrre venti. Nella bassa ionosfera, tuttavia, i protoni collidono troppo frequentemente con le particelle neutre e vengono espulsi dalla deriva. Perciò solo gli elettroni possono fluire lungo l'ovale aurorale: la corrente elettrica netta è diretta verso est nel settore della sera e verso ovest in quello del mattino.

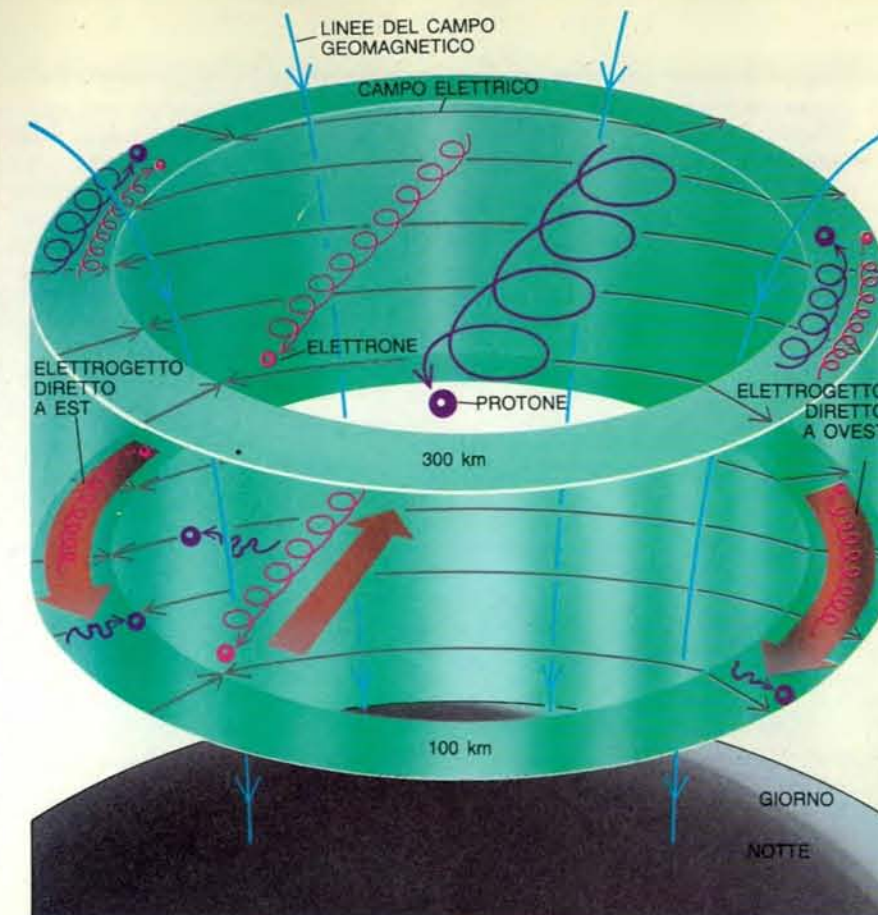
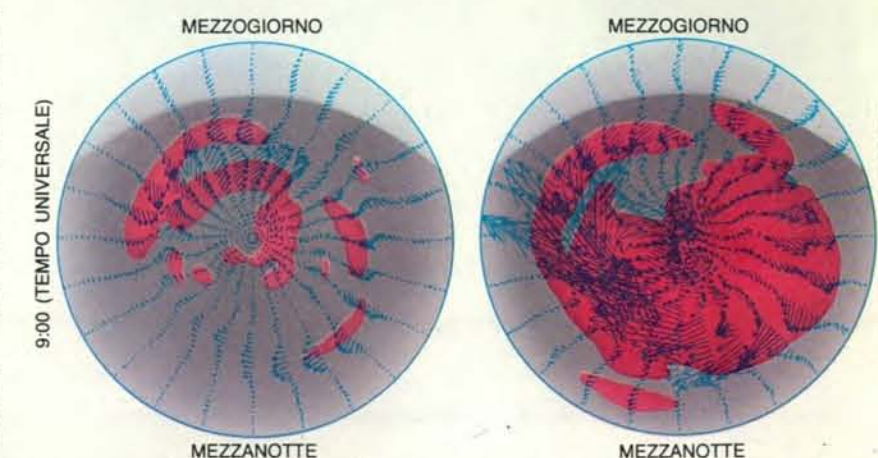
Durante l'Anno geofisico internazionale (1957-1958) le osservazioni fotografiche dell'intera volta celeste rivelarono un aspetto caratteristico, fino a quel momento sconosciuto, dell'attività aurorale sopra le regioni polari: la cosiddetta sottotempesta aurorale. La prima indicazione di una sottotempesta è un'improvvisa accentuazione della luminosità del drappeggio aurorale che avviene tra la sera inoltrata e mezzanotte. Questo fenomeno si estende rapidamente lungo il drappeggio in entrambe le direzioni, cosicché dopo alcuni minuti l'intera parte del drappeggio che si trova nell'emisfero notturno diventa luminosa.

Il drappeggio brillante comincia a muoversi verso il polo nel settore della mezzanotte alla velocità di alcune centinaia di metri al secondo. Contemporaneamente qui si osserva la formazione di un rigonfiamento di grande scala. Diversi movimenti ondulatori appaiono nei pressi dell'estremità occidentale del rigonfiamento e si estendono verso ovest alla velocità di un chilometro al secondo. A metà strada intorno al polo, nel settore del mattino, i drappeggi aurorali si scompongono in molte «chiazze». Il moto verso il polo nel settore della mezzanotte dura in genere da 30 minuti a un'ora; dopo che questo avanzamento ha raggiunto la latitudine massima l'attività aurorale comincia a ridursi. Una sottotempesta dura in genere da una a tre ore.

La sottotempesta aurorale è una manifestazione di una sottotempesta magnetosferica, che avviene circa quattro o cinque volte al giorno. Tra le altre manifestazioni vi è la formazione di elettrogetti molto intensi, che a loro volta causano forti perturbazioni geomagnetiche, le sottotempeste magnetiche polari. Una di queste tempeste, avvenuta il 18 marzo 1978, è stata registrata nei particolari grazie a una collaborazione internazionale che ha permesso di collocare più di 70 magnetometri nella regione artica, disposti lungo sei «raggi» centrati sul polo nord magnetico. Grazie a raffinate elaborazioni al calcolatore, Yosuke Kamide della Kyoto Sangyo University e Yasha I. Feldstein dell'Accademia delle scienze dell'Unione Sovietica sono stati in grado di ricostruire dalle registrazioni magnetiche l'andamento delle correnti elettriche. Hanno anche potuto stimare il tasso di produzione di calore associato con queste correnti ionosferiche.

La dinamica della sottotempesta

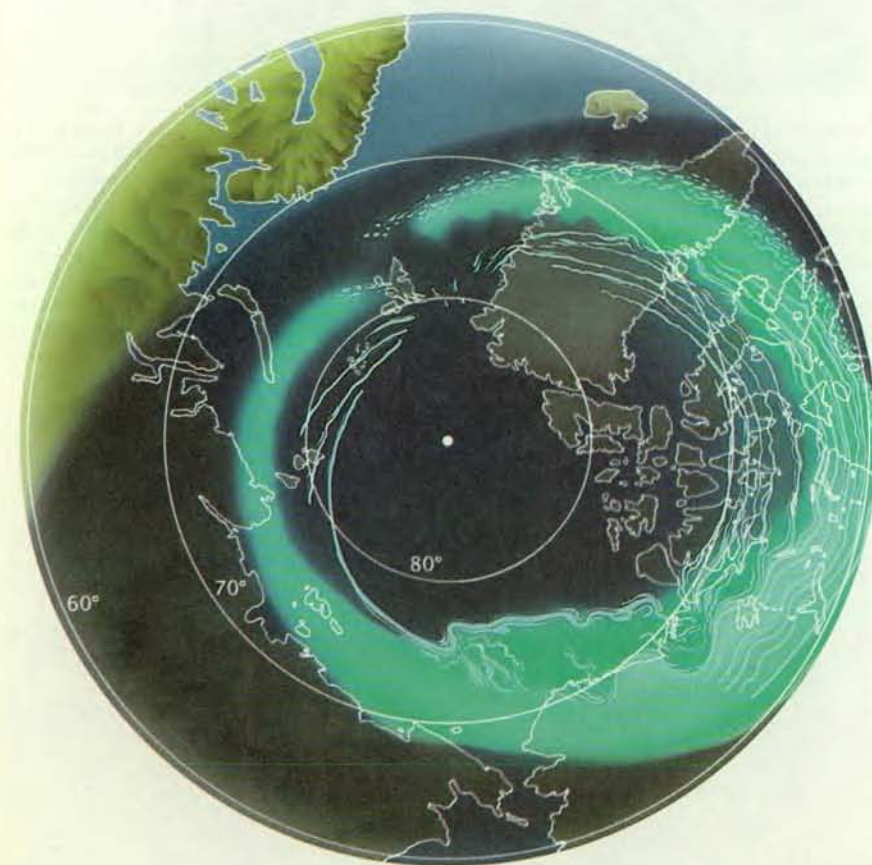
Secondo molti ricercatori, fra cui Robert McPherron dell'Università della California a Los Angeles e Daniel Baker del Goddard Space Flight Center della National Aeronautics and Space Administration, le sottotempeste sarebbero provocate da processi che avvengono nella coda magnetosferica. Il processo del generatore produce due correnti ad anello che fluiscono in direzioni opposte nei lobi settentrionale e meridionale della coda magnetosferica. Le correnti circolanti inducono campi magnetici antiparalleli lungo tutta la coda. Si ritiene generalmente che questi campi antiparalleli si riconnettano spontaneamente in maniera esplosiva, liberando l'energia che alimenta le sottotempeste.

Tuttavia è diventato sempre più chiaro come la crescita e l'attenuazione delle sottotempeste magnetosferiche siano in parte controllate dall'aumento e dalla diminuzione dell'energia prodotta dal generatore. All'inizio degli anni settanta Paul Perreault, un mio laureando, e io abbiamo tentato di correlare l'energia generata da questo processo con varie caratteristiche del vento solare. Abbiamo supposto che l'effettiva energia totale misurata che viene dissipata nella ma-

Le correnti allineate con il campo inducono nell'ovale aurorale campi elettrici perpendicolari al campo geomagnetico. In questa situazione gli elettroni e i protoni sono deviati nella stessa direzione, un fenomeno noto come deriva E vettor B (E rappresenta il campo elettrico, B il campo induzione magnetica). La deriva sulla calotta polare ha direzione opposta al moto lungo l'ovale dato che i campi elettrici puntano in direzioni opposte. (Le particelle seguono traiettorie a spirale perché la forza di Lorenz imprime loro un moto circolare intorno alle linee di forza del campo magnetico.) Nell'alta ionosfera elettroni e protoni si muovono alla stessa velocità e non vi è corrente netta. Nella bassa ionosfera invece i protoni subiscono molte collisioni e si muovono di poco in direzione del campo elettrico, ma non in quella della deriva. Si ha quindi una corrente netta di elettroni, che produce gli elettrogetti.

Le correnti ionosferiche registrate il 18 marzo 1978 durante una sottotempesta (*a destra*) sono molto più estese di quelle rilevate appena un'ora prima (*a sinistra*). Il polo nord magnetico si trova al centro dell'illustrazione, con il lato diurno del pianeta in alto e quello notturno in basso. Le frecce puntano nella direzione delle correnti elettriche e la loro lunghezza ne indica l'intensità. Durante la sottotempesta sono stati osservati un elettrogetto diretto verso ovest nel settore notturno e uno diretto verso est nel settore del pomeriggio. L'intensità del colore rosso indica la quantità di calore generato dall'attività elettrica.

gnetosfera interna sia uguale al tasso di immissione di energia nella magnetosfera da parte del vento solare. Abbiamo poi tentato di vedere se le fluttuazioni dell'energia dissipata siano correlate con le variazioni di talune caratteristiche del vento solare misurate dai satelliti.

Abbiamo determinato che l'energia è proporzionale al prodotto della velocità del vento solare, del quadrato dell'intensità del campo magnetico e della quarta potenza del seno di metà dell'angolo polare (misurato a partire dal Polo nord) secondo il quale questo campo incide sul campo magnetico terrestre. In altri termini, l'energia è nulla quando il campo magnetico del vento solare è diretto verso nord, perché in questo caso l'angolo polare è zero e così pure la sua funzione seno. Al contrario l'energia è massima quando il campo è diretto verso sud (a parità di altre condizioni) perché l'angolo polare è di 180 gradi e la funzione seno assume il valore massimo.

Mikhail Pudovkin e colleghi dell'Università di Leningrado, insieme con altri fisici, hanno ricavato teoricamente la stessa formula supponendo che la magnetosfera si comporti come un generatore magnetoidrodinamico. Patricia H. Reiff e colleghi sono riusciti a dimostrare che le variazioni misurate della caduta di potenziale sulla calotta polare (la tensione prodotta dal generatore aurorale) sono strettamente correlate all'energia calcolata tramite la nostra equazione. Per di più, ogni significativo aumento dell'energia al di sopra di 10 000 megawatt è associato non solo a un incremento della caduta di potenziale, ma anche a un'intensificazione delle sottotempeste aurorali.

Si osserva che l'ovale aurorale si contrae e si dilata in relazione alla quantità di energia fornita dal generatore aurorale, la quale a sua volta è funzione della componente nord-sud del campo magnetico interplanetario. Queste osservazioni indicano che nella maggior parte dei casi le sottotempeste avvengono quando il vettore campo magnetico del vento solare è diretto verso sud; ciò dimostra in modo convincente che le sottotempeste sono controllate direttamente dal vento solare e non da eventi spontanei entro la magnetosfera.

Un particolare dell'ovale aurorale è mostrato in questa immagine che ha il Polo nord geografico al centro. Un'aurora diffusa si estende in una fascia ininterrotta dal settore della sera a quello della mezzanotte. Nel settore della sera (*a sinistra*), ha una luminosità abbastanza uniforme ed è ampia a sufficienza da riempire l'intero campo visivo di un osservatore a terra. Aurore separate, a forma di drappeggio, sono distribuite lungo il lato dell'aurora volto verso il polo. Al culmine di una sottotempesta forme luminose avanzano in direzione del polo intorno alla mezzanotte. Verso il settore del mattino la porzione diffusa sembra disgregarsi in strutture a forma di drappeggio e in chiazze sul perimetro esterno dell'ovale.

Alcuni problemi irrisolti

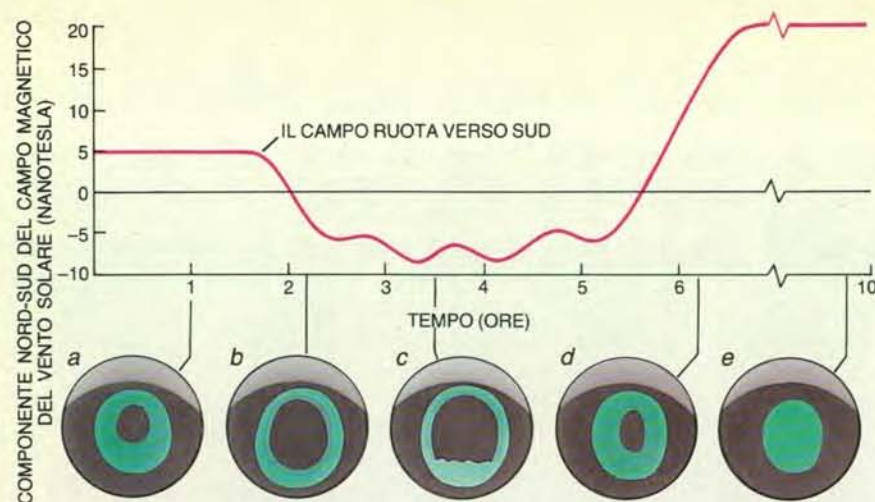
Non si sa ancora esattamente che cosa inneschi una sottotempesta. In particolare, un aumento dell'energia prodotta dal generatore aurorale non può spiegare di per sé l'improvvisa accentuazione della luminosità dei drappeggi. Secondo Joseph R. Kan dell'Università dell'Alaska a Fairbanks questo fenomeno sarebbe provocato dall'amplificazione delle correnti allineate con il campo, dovuta a effetti che insorgono nella ionosfera quando il processo della deriva **E** vettor **B** viene accentuato da un incremento dell'energia prodotta.

Che cosa accade quando il campo magnetico è intenso e diretto verso nord per un lungo periodo, anziché verso sud? Via via che l'energia del generatore aurorale diminuisce, l'aurora diventa sempre più pallida e gli elettrogetti più deboli. In queste condizioni avviene un fenomeno aurorale inatteso. Diversi drappeggi aurorali si estendono sulla calotta polare in direzione parallela al meridiano mezzogiorno-mezzanotte e aurore a chiazze, non visibili a occhio nudo, si spostano sulla calotta polare nella stessa direzione. Non è possibile spiegare queste aurore semplicemente in termini di una diminuzione dell'energia del generatore e la loro trattazione non rientra nell'argomento di questo articolo.

Il modello della produzione di energia aurorale di cui ho parlato finora riguarda la riconnessione delle linee di forza del campo magnetico sul lato della magnetosfera rivolto verso il Sole. Si suppone che la riconnessione abbia luogo anche nella coda magnetosferica, mediante un processo che risulta anch'esso indotto e in parte controllato dal generatore aurorale. Secondo dati raccolti da James A. Slavin e Bruce T. Tsurutani del Jet Propulsion Laboratory, la velocità del plasma in allontanamento dalla Terra, a una distanza a valle di circa 200 raggi terrestri, appare correlata all'intensità degli elettrogetti aurorali. Si ritiene che il flusso a valle venga prodotto dall'energia liberata dalla riconnessione delle linee di forza nella coda magnetosferica.

Via via che l'aurora diventa più luminosa, gli strati di plasma della coda magnetosferica si fanno più sottili e in breve tempo danno origine a molteplici forme di attività nella coda magnetosferica. Si ritiene che queste ultime siano in qualche modo correlate alla dinamica dell'aurora durante una sottotempesta. Il progetto International Solar-Terrestrial Physics, un vasto programma di ricerca con numerosi satelliti attualmente allo stadio finale di programmazione, dovrebbe essere in grado di rispondere a questi interrogativi.

Oltre le aurore

L'energia del generatore aurorale è controllata in definitiva dall'attività solare, in particolare da eventi transienti

Quando il vettore campo magnetico del vento solare punta verso nord l'ovale aurorale è piccolo e pervaso da un bagliore non visibile a occhio nudo (*a*). Via via che il campo ruota verso sud l'ovale diventa più luminoso e si allarga rapidamente; il bagliore scompare, tranne che lungo una stretta fascia all'interno dell'ovale (*b*). Circa un'ora più tardi inizia una sottotempesta e drappeggi luminosi avanzano verso il polo. La sottotempesta raggiunge il culmine una o due ore dopo il suo inizio (*c*). Dopo che il vettore campo magnetico si volge di nuovo verso nord, l'aurora impallidisce e drappeggi paralleli al meridiano mezzogiorno-mezzanotte appaiono sulla calotta polare (*d*). Quando il vettore campo magnetico possiede una grande componente diretta verso nord per molte ore, il bordo interno dell'ovale talvolta svanisce lasciando un bagliore uniforme su tutta la regione polare (*e*).

quali i brillamenti solari e l'espulsione di grandi quantità di gas coronale, che generano onde d'urto propagantisi nel vento solare. Subito dopo il fronte d'onda il vento solare raggiunge una velocità compresa fra 500 e 1000 chilometri al secondo, cosicché il campo magnetico interplanetario viene compresso e perciò incrementato. Quando l'onda d'urto collide con la magnetosfera, l'energia prodotta dal generatore può balzare a 10 milioni di megawatt o più, purché il campo magnetico sia diretto verso sud.

In questa situazione può prodursi una tempesta geomagnetica, durante la quale l'ovale aurorale si allarga in modo anomalo. Nella sua espansione verso sud esso può svanire dal cielo dell'Alaska e apparire presso il confine tra Stati Uniti e Canada o anche a latitudini più basse. L'emissione rossa degli atomi di ossigeno, alla lunghezza d'onda di 630 nanometri, risulta notevolmente accentuata in queste aurore, forse perché l'incremento di energia eccita termicamente gli atomi di ossigeno a livelli energetici più elevati. Contemporaneamente l'intensificazione delle correnti nelle fasce di Van Allen dà origine a intensi campi magnetici anche alle basse latitudini e al suolo.

Recentemente i miei colleghi e io abbiamo tentato di comprendere l'effetto delle onde d'urto solari sulla magnetosfera. Un brillamento al centro del disco solare genera un'onda d'urto che si propaga verso la Terra lungo la congiungente Terra-Sole. In questa situazione il fronte d'onda collide quasi direttamente con il fronte della magnetosfera; il campo magnetico interplanetario subisce una notevole compressione e quindi la sua intensità aumenta. A sua volta questo significa che anche l'energia prodotta dal generatore aurorale risulta molto maggiore. Se il brillamento avviene presso il bordo del disco solare, l'onda d'urto si propaga in direzione perpendicolare alla congiungente Terra-Sole, sfiorando solamente il fronte della magnetosfera, e quindi la comprime in misura insignificante. In questa situazione è possibile che anche un brillamento intenso non dia luogo a manifestazioni aurorali spettacolari.

Un altro fenomeno solare che influenza l'attività aurorale è rappresentato dai «buchi» coronali (regioni prive di macchie solari). Questi buchi, che sono particolarmente sviluppati durante la fase declinante del ciclo di attività solare, generano flussi di vento solare di alta velocità. Spesso compaiono simultaneamente due buchi di grandi dimensioni, uno dei quali si estende a partire dal polo nord solare e l'altro dal polo sud; ognuno di essi espelle vento solare in un ampio getto.

Dato che il Sole ruota su se stesso con un periodo di circa 27 giorni rispetto alla Terra, vi è un effetto simile a quello prodotto da un annaffiatore rotante: un flusso arriva sulla Terra seguito a distanza di due settimane dall'altro. La Terra rimane immersa in ciascun flusso per circa una settimana e durante questo periodo l'energia prodotta dal generatore aurorale è elevata e variabile. I buchi coronali tendono a durare per molti mesi, perciò nella fase declinante di un ciclo di attività solare vi saranno due periodi di attività aurorale, ciascuno della durata di una settimana, ogni 27 giorni, e questo per parecchi mesi. Sembra che il vento solare venga emesso a una velocità maggiore alle latitudini più elevate di un buco coronale; questo fatto potrebbe spiegare per quale motivo l'attività aurorale si intensifica nei mesi in cui cadono gli equinozi di primavera e di autunno, quando la Terra si trova alle più elevate latitudini eliocentriche.

L'elemento cruciale per comprendere taluni fenomeni solari potrebbe in effetti trovarsi nelle aurore stesse. Come le aurore, i brillamenti solari sono provocati dalle emissioni di atomi eccitati nell'atmosfera e appaiono in forma di strutture luminose simili a drappeggi; senza dubbio sono manifestazioni di processi simili. Da parecchio tempo si suppone che l'energia dei brillamenti sia fornita dalla riconnessione delle linee di forza del campo magnetico del Sole. La teoria richiede l'esistenza dei cosiddetti campi liberi da forze, che sono identici a quelli associati alle correnti allineate con il campo nella magnetosfera terrestre. È essenziale scoprire un meccanismo generatore che possa fornire l'energia elettrica necessaria per la formazione dei campi liberi da forze e dei brillamenti solari. Forse i moti dei gas sulla superficie visibile del Sole, che è simile alla ionosfera terrestre, sono in grado di produrre l'energia richiesta.

Confido che lo studio delle aurore possa contribuire alla conoscenza di un gran numero di fenomeni astrofisici. In definitiva, plasmi rarefatti permeati da campi magnetici fanno parte di quasi tutti gli oggetti astronomici; le interazioni tra i flussi di plasma magnetizzato e l'atmosfera di corpi celesti dotati di campo magnetico, come stelle, pianeti e comete, potrebbero essere molto comuni. Di tutte queste interazioni solo l'aurora fornisce un laboratorio naturale nel quale gli studiosi possono verificare le loro teorie mediante l'osservazione diretta e trovarvi conferma.

BIBLIOGRAFIA

AKASOFU SYUN-ICHI e CHAPMAN SYDNEY, *Solar-Terrestrial Physics*, Oxford University Press, 1972.

JONES A. VALLANCE, *Aurora*, D. Reidel Publishing Co., 1974.

EATHER ROBERT H., *Majestic Lights*, American Geophysical Union, 1980.

BREKKE ASGEIR ed EGELAND ALF, *The Northern Lights: From Mythology to Space Research*, Springer-Verlag, 1983.

AKASOFU SYUN-ICHI e KAMIDE YOSUKE, *The Solar Wind and the Earth*, D. Reidel Publishing Co., 1987.

Passato e futuro dell'Amazzonia

La storia climatica della foresta pluviale amazzonica ci insegna che questo ecosistema è in grado di reagire positivamente alle perturbazioni naturali, ma non certo all'attuale sfruttamento indiscriminato da parte dell'uomo

di Paul A. Colinvaux

La foresta pluviale che copre il vasto bacino del Rio delle Amazzoni sembra dall'alto un uniforme tappeto verde solcato qua e là da corsi d'acqua. In realtà, la foresta è tutto fuorché uniforme: quello che ci appare come un tappeto è il baldacchino della foresta, formato dall'ampio fogliame di diversi generi di alberi giganteschi e, a sua volta, questo baldacchino non è che lo strato più alto di un ecosistema che ospita un numero di specie superiore a quello di qualsiasi altra regione della Terra. Nella foresta pluviale vivono probabilmente 80 000 specie vegetali (tra cui 600 varietà di palme) ed è verosimile che essa accolga 30 milioni di specie animali, per la maggior parte insetti.

Un tempo si riteneva che questa straordinaria ricchezza di forme di vita fosse il risultato di una situazione climatica stabilizzatasi da moltissimo tempo, caratterizzata da alte temperature e piogge abbondanti. Secondo questa teoria il dominio tropicale conosciuto semplicemente come Amazzonia, risparmiato dalla «catastrofe» a breve termine rappresentata ogni anno dall'inverno e da quelle a lungo termine costituite dalle glaciazioni, avrebbe perso poche specie e ne avrebbe al contrario accumulate molte. Tuttavia i dati di campagna inducono a ritenere che l'Amazzonia sia soggetta a cambiamenti climatici su tutte le scale temporali, compreso un raffreddamento verificatosi quando le condizioni climatiche dell'età glaciale provocarono l'espansione dei ghiacciai su gran parte dell'emisfero settentrionale. Inoltre, le moderate perturbazioni climatiche della regione, anziché danneggiare le forme di vita, possono in realtà fornire una spiegazione per la splendida diversificazione dell'attuale foresta pluviale amazzonica.

Lo studio di come le condizioni climatiche del passato possano aver influito sulla diversità di specie dell'ecosistema amazzonico va al di là del semplice interesse accademico in quanto può servire da strumento di previsione della capacità della foresta di sopportare i cambiamenti. A mano a mano che gli esseri umani procedono nella devastazione di ampie zone di foresta, un numero inusitato e incalcolabile di specie si va estinguendo. Questo impone alle nazioni sotto la cui giurisdizione si trova la foresta pluviale, Brasile, Venezuela, Colombia, Ecuador e Perù, di individuare strategie di gestione che, pur compatibili con le loro esigenze di sviluppo economico, siano in grado di salvaguardare il maggior numero possibile di specie. Coloro che cercano di mettere a punto tali strategie devono conoscere i fattori che possono danneggiare la diversità delle specie, ragion per cui è importante scoprire quali tipi di aggressione abbia sopportato l'ecosistema nel corso della sua storia.

La vecchia ipotesi secondo cui la varietà di specie riscontrabile in Amazzonia sarebbe dipesa dalla stabilità climatica della regione era basata, abbastanza curiosamente, su osservazioni effettuate in mare profondo. Howard L. Sanders del Marine Biological Laboratory di Woods Hole aveva infatti notato una grande varietà di specie fra gli animali che dimorano nel fango dei fondali oceanici profondi, nonostante la fredda oscurità e la bassa produttività biologica di quell'ambiente; ne aveva dedotto che le condizioni inospitali fossero compensate dalla stabilità del luogo. In assenza di significative variazioni nelle condizioni fisiche, l'estinzione di specie adattate alla situazione dominante dovrebbe essere un evento raro. Nel corso del tempo, nuove specie continuerebbero a evolversi, cosicché la velocità di speciazione supererebbe quella di estinzione con un incremento nella diversità delle specie.

Sanders propose di considerare la foresta dell'Amazzonia e altre foreste tropicali come ambienti analoghi (anche se più produttivi) a quello di mare profondo, cioè come luoghi dove, per la costanza di umidità e temperatura durante l'anno, l'estinzione sarebbe rara. A sostegno di questa ipotesi c'erano l'assenza di una stagione invernale e di glaciazioni, nonché reperti in base ai quali si poteva supporre che specie vegetali della foresta pluviale fossero sopravvissute per 30 milioni di anni o più in qualche luogo del bacino del Rio delle Amazzoni.

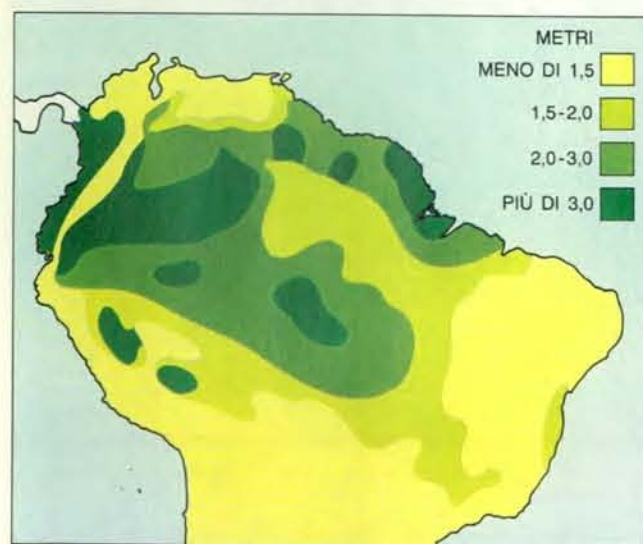
Successivamente, nel 1969, diverse osservazioni gettarono alcuni dubbi sulla validità della teoria della stabilità dell'Amazzonia, inducendo a pensare che le condizioni climatiche in questa regione avessero subito fluttuazioni significative nel passato. Jürgen Haffer, allora ricercatore alla Mobil Research and Development Corporation, rilevò che zone ristrette del bacino, dalle caratteristiche peculiari, possedevano una loro specifica fauna di uccelli nonostante la foresta verde si estendesse praticamente ininterrotta dal margine occidentale dell'Amazzonia fino all'Atlantico. Haffer e altri biologi rilevarono anche altre divergenze specifiche nelle farfalle e in altri gruppi animali. Questo fenomeno costituiva un interessantissimo rompicapo per i biogeografi: per quale motivo alcune popolazioni sarebbero rimaste isolate se il loro habitat era continuo?

Haffer propose una spiegazione talmente chiara e logica che molti di noi si rammaricarono di non averci pensato. Secondo il ricercatore, l'attuale isolamento affonda le sue radici nell'ultima età glaciale. Osservando che le regioni in cui si verificano fenomeni di isolamento specifico hanno il proprio centro su rilievi e che i terreni pianeggianti sono più asciutti di quelli montagnosi, egli propose un modello in base al quale nelle età glaciali i bassopiani dell'Amazzonia sarebbero diventati una distesa di terreni semidesertici, mentre le regioni più elevate si sarebbero mantenute umide fungendo quindi da isolati rifugi per la flora e la fauna della foresta pluviale. Popolazioni con una distribuzione precedentemente continua si diversificarono lungo questo gradiente ecologico per separarsi poi definitivamente adattandosi ai nuovi habitat in quota.

L'ipotesi è affascinante perché sembra spiegare non solo la divergenza tra le specie che vivono in Amazzonia, ma anche la loro inconsueta varietà. I rifugi dell'età glaciale, come prevedeva la precedente ipotesi della stabilità, avrebbero protetto le specie esistenti dall'estinzione, ma il periodico isolamento geografico di popolazioni imparentate (secondo diverse stime ci sarebbero state 13 età glaciali fino a oggi) avrebbe facilitato una divergenza sempre crescente.

Perché l'ipotesi dei rifugi regga, si dovrebbe riuscire a dimostrare che i bassopiani dell'Amazzonia, su cui ora si estende la foresta pluviale, divennero aridi durante le età glaciali. Nessuno finora ha trovato prove inconfutabili che l'Amazzonia fosse effettivamente arida, sebbene una testimonianza indiretta sia stata fornita da elaborazioni al calcolatore e da campioni di sedimenti provenienti da numerosi siti tropicali.

Vista dall'alto, quest'area di foresta pluviale amazzonica solcata da corsi d'acqua appare uniforme, ma, come il resto della foresta, comprende in realtà molte specie di alberi che crescono fianco a fianco. In effetti questa foresta, calda e umida, ospita un numero di specie animali e vegetali per unità di superficie (e anche in totale) maggiore di qualsiasi altra regione della Terra. Un tempo si pensava che questa varietà derivasse da condizioni climatiche costanti, ma ora sembra che in Amazzonia non siano mancate tempeste locali e inondazioni. A differenza delle glaciazioni che provocano estinzioni in massa, queste perturbazioni più moderate non comportano estinzioni e, anzi, eliminando qualche individuo delle specie dominanti, offrono a quelle più deboli le opportunità per affermarsi.

Il bacino del Rio delle Amazzoni è delimitato - tranne la regione affacciata sull'Atlantico - da alte catene montuose: gli altopiani del Venezuela e della Guiana a nord, le Ande a ovest e il Mato Grosso a sud. La foresta pluviale si estende fino a 1200 metri di quota.

Le mappe delle medie pluviometriche annue (*a sinistra*) e degli habitat di diverse specie di farfalle (*a destra*) mostrano che le regioni in cui si ha l'isolamento delle specie si sovrappongono in genere a quelle più umide; esse inoltre hanno il centro su colline (non indicate). Tale distribuzione, notata dapprima per gli uccelli, suggerì l'ipotesi che gli habitat separati si fossero instaurati molto tempo fa, forse nell'ultima età glaciale (da 70 000 a 10 000 anni fa): nelle età glaciali i bassopiani diverrebbero aridi e i rilievi, più umidi, si trasformerebbero in rifugi per le specie tipiche della foresta pluviale; evolutesi separatamente, le specie rimarrebbero isolate anche quando la foresta si rigenera. Nuove scoperte indicano che i rifugi ipotizzati erano troppo freddi perché vi potesse sopravvivere la foresta pluviale. Questa sarebbe persistita in qualche area dei più caldi bassopiani, che non si sarebbero inariditi. Pur escludendo la teoria dei rifugi, tale conclusione avvalora il presupposto che l'Amazzonia abbia subito sensibili fluttuazioni climatiche.

Per esempio, John E. Kutzbach e Peter J. Guetter dell'Università del Wisconsin a Madison, simulando al calcolatore il clima del globo nelle età glaciali, hanno sviluppato modelli in base ai quali le piogge monsoniche ai tropici risulterebbero ridotte del 20 per cento. Una riduzione di questa entità avrebbe sicuramente fatto estendere l'aridità nelle regioni con una forte caratterizzazione stagionale, dove cioè annualmente si hanno periodi molto secchi, regioni fra le quali rientra per lo meno qualche zona dell'Amazzonia. (Molte delle regioni tropicali soggette a variazioni stagionali non hanno veri periodi di clima secco, ma solo stagioni meno piovose di altre.)

A Daniel A. Livingstone della Duke University si devono alcuni dei primi dati di campagna in grado di fornire qualche chiarimento sul problema dell'aridità. Esaminando polline fossile di alberi, un buon indicatore della flora locale, il ricercatore scoprì che diverse foreste pluviali tropicali dell'Africa attuale sorgono dove, tra 20 000 e 12 000 anni fa, cioè verso la fine dell'ultima età glaciale (durata da 70 000 a 10 000 anni fa circa), si estendeva la savana. Accenni a una simile aridità verificatasi nell'età glaciale in quelle che sono attualmente regioni tropicali umide giunsero anche dalla periferia del bacino del Rio delle Amazzoni, suggerendo che quanto era stato accertato per l'Africa dell'età glaciale poteva esser avvenuto anche nel bacino del Rio delle Amazzoni.

Io stesso ho potuto raccogliere dati a sostegno dell'ipotesi di un'Amazzonia arida alle isole Galápagos, che si trovano all'incirca 2000 chilometri a ovest rispetto alla estremità occidentale del bacino del Rio delle Amazzoni, in corrispondenza dell'equatore. Una carota di sedimenti ben databile, prelevata dal fondale di un lago di acqua dolce delle Galápagos, mostra che il bacino era asciutto durante l'ultima età glaciale e non si riempì prima dell'inizio dell'Olocene (l'attuale periodo postglaciale).

Sul continente, altri hanno trovato testimonianze simili in siti a nord dell'Amazzonia. Nel Venezuela settentrionale, la regione oggi occupata dal profondo lago di Valencia era priva di acqua alla fine dell'ultima età glaciale e le zone costiere attualmente palustri della Guyana presentavano una vegetazione caratteristica di un clima secco. Diverse ricerche suggeriscono che fosse arida anche la regione a sud del bacino amazzonico, presso São Paulo, dove ora sorge una foresta con caratteristiche marcatamente stagionali.

Ulteriori testimonianze del fatto che i bassopiani dell'Amazzonia durante l'età glaciale potessero essere aridi sono state ottenute da carote prelevate in alto mare nell'Oceano Atlantico, proprio nei pressi della foce del Rio delle Amazzoni. John E. Damuth e Rhodes W. Fairbridge, allora ricercatori presso la Columbia University, hanno trovato che i sedimenti trasportati verso est dal fiume nel corso dell'ultima età glaciale includevano una piccola frazione di feldspati, minerali provenienti da rocce cristalline, che potevano plausibilmente essere prodotti dall'erosione di una regione arida.

Nonostante l'abbondanza delle prove indirette a favore di una fase arida durante l'età glaciale, mi sono convinto che non si tratta di testimonianze incontrovertibili. Georg Irion dell'Istituto di ricerca e Museo di storia naturale di Francoforte ha dimostrato che i feldspati presenti nei depositi atlantici provenivano probabilmente da rocce del bacino amazzonico, ma non da rocce superficiali. Durante l'età glaciale il livello del mare si abbassò e molti corsi d'acqua probabilmente approfondirono i loro alvei, asportando i feldspati nel corso di questo processo.

Anche i dati ottenuti dalle carote di sedimenti provenienti dalle isole Galápagos e dalle regioni sudamericane che circondano l'Amazzonia sono poco decisivi, perché le zone di provenienza dei sedimenti, benché più vicine dell'Africa, sono ancora troppo distanti dalla foresta pluviale amazzonica. Inoltre tutte queste regioni sono spesso soggette all'influenza di fattori climatici differenti. Tra i luoghi studiati e il bacino del Rio delle Amazzoni si trovano i picchi delle Ande, gli altopiani del Venezuela e della Guiana e l'immensa linea di rilievi del Mato Grosso (grande boscaglia), che costituiscono rispettivamente i confini occidentale, settentrionale e meridionale del bacino. Ciascuno di questi confini è sede di cambiamenti climatici e di conseguenza il clima è spesso diverso da un versante all'altro.

È possibile che nelle regioni della attuale Amazzonia con caratteristiche marcatamente stagionali la riduzione delle precipitazioni monsoniche abbia dato luogo a semiaridità nelle età glaciali, forse proprio ai piedi degli altopiani della Guiana nell'angolo nordorientale, immediatamente a ovest del Mato Grosso nella regione sudoccidentale e in alcune parti del bacino centrale. Per gran parte delle regioni più umide dell'Amazzonia, in cui le precipitazioni annuali variano fra i 2000 e i 5000 millimetri, è abbastanza improbabile che anche una riduzione di piovosità del 20 per cento produca qualche differenza rilevante per le forme di vita.

Nel 1985 la totale mancanza di testimonianze fossili attribuibili con il metodo del radiocarbonio all'età glaciale e provenienti dai rifugi ipotizzati da Haffer o dai bassopiani amazzonici indusse me e alcuni colleghi della Ohio State University a compiere ricerche nell'Amazzonia vera e propria. Trovammo i soli fossili che a tutt'oggi siano stati attribuiti all'età glaciale. Essi indicavano che l'Amazzonia non era arida, ma ave-

***Podocarpus oleifolius* è solo una delle diverse specie di *Podocarpus* che crescono intorno ai 1800 metri (e più) nelle Ande ecuadoriane. Queste piante, dal momento che richiedono le condizioni relativamente fresche delle elevate altitudini, non vivono nella corrispondente parte della foresta pluviale amazzonica. L'autore ha trovato recentemente nella foresta pluviale dell'Ecuador reperti fossili di *Podocarpus* risalenti all'età glaciale. È in primo luogo questa scoperta a suggerire che la teoria dei rifugi per spiegare la divergenza specifica non sia corretta: essa infatti indica che il sito di ritrovamento dei fossili, che si trova all'interno di uno dei principali rifugi ipotizzati, a Mera, era troppo freddo per permettere la sopravvivenza della foresta pluviale. Se ne deve quindi dedurre che probabilmente anche alcuni degli altri rifugi ipotizzati (o tutti) si trovavano nelle medesime condizioni.**

Il lago Añangucocha è uno dei quattro laghi ecuadoriani i cui reperti sedimentari testimoniano che la parte occidentale della foresta pluviale amazzonica subì un lungo periodo caratterizzato da tempeste. Lo strato superficiale di sedimento è composto di *gyttja*, tipico fango lacustre; sotto di esso vi è uno strato di caratteristico sedimento di fiume, che venne depositato sopra un altro strato di fango lacustre tra gli 800 e i 1300 anni fa. Questa distribuzione mostra che circa 1300 anni fa nell'Añangucocha, che in precedenza era rimasto isolato, si reintrodusse il suo immissario, riempiendolo per circa 500 anni, probabilmente perché il fiume si era ingrossato grazie a ripetuti e potenti uragani. Il fiume abbandonò il lago circa 800 anni fa, verosimilmente quando le tempeste vennero meno.

va comunque subito un cambiamento climatico significativo: una diminuzione di temperatura di vari gradi Celsius.

Trovammo i nostri esemplari per caso, mentre esploravamo la regione al margine più settentrionale della foresta pluviale, nell'Ecuador orientale, alla ricerca di laghi nei cui sedimenti contavamo di trovare testimonianze sui climi remoti. A un certo punto ci trovammo su una strada sterrata a Mera che, snodandosi a un'altitudine di 1100 metri nei rilievi pedemontani orientali delle Ande, corre in prossimità del limite superiore della foresta pluviale attuale e all'interno di uno dei principali ipotetici rifugi. Qui notammo antichi tronchi e ceppi rimasti inclusi in uno strato di sedimenti affioranti. Approfittammo del tempo di attesa dell'aereo per raccogliere campioni di sedimenti e legno. La datazione del legno con il radiocarbonio dimostrò che alcuni campioni risalivano a 35 000 anni fa e altri a 26 000 anni fa. Per quanto anteriori al culmine della fase più fredda dell'ultima età glaciale (circa 18 000 anni fa) queste datazioni si riferiscono comunque a età glaciali.

Altre analisi dei campioni di legno hanno dimostrato che questi comprendevano legno tenero, di conifere. Attualmente le uniche conifere presenti nell'Ecuador appartengono al genere *Podocarpus*, tipico delle foreste andine, che oggi cresce almeno 700 metri più in alto rispetto alla quota di Mera. In effetti, analisi polliniche dei campioni di sedimenti hanno evidenziato che l'antica foresta di Mera aveva marcate caratteristiche andine e che di essa non esiste un diretto corrispettivo nell'attuale America tropicale.

Possiamo concludere che gli alberi del genere *Podocarpus*, che richiedono il clima umido e relativamente fresco oggi riscontrabile solo sulle montagne, nelle età glaciali crescevano almeno 700 metri più in basso. Questa considerazione, unitamente ai dati ottenuti dall'esame dei pollini, ci suggeriva che nell'ultima età glaciale Mera fosse piovosa, ma troppo fredda per ospitare un tipo di foresta pluviale simile all'attuale. Secondo una formula standard, la temperatura sui rilievi dell'Ecuador doveva essere almeno di 4,5 gradi Celsius più bassa rispetto all'attuale. (Secondo questa formula nell'aria umida la temperatura diminuisce, lungo il versante di una montagna, di sei gradi Celsius ogni 1000 metri.)

Inoltre, Mera non offriva alcun rifugio alle specie termofile. Se l'abbassamento della temperatura avesse riguardato anche l'Amazzonia, come sembra probabile, altre parti dell'odierna foresta pluviale situate in quota potrebbero aver raggiunto temperature troppo basse per costituire rifugi per la foresta pluviale. Dal momento che alberi tipici della foresta pluviale dovettero sopravvivere in qualche luogo del bacino amazzonico, i dati suggeriscono che la foresta pluviale si fosse ritirata in qualche zona delle regioni a quote inferiori, laddove secondo i teorici dei rifugi avrebbe dovuto esservi aridità. Se effettivamente gli alberi sopravvissero solo nei bassopiani, va da sé che questi non potevano essere aridi, almeno non uniformemente.

Ovviamente, per ricostruire in modo attendibile le condizioni dell'Amazzonia durante l'età glaciale occorre una maggiore quantità di dati risalenti a quell'epoca, particolarmente alle fasi più fredde di quel periodo. A tutt'oggi, alla luce delle prove esistenti, l'ipotesi dei rifugi non sembra più sostenibile dell'ipotesi che ha rimpiazzato, quella della persistenza di condizioni uniformi. Anzi, i dati suggeriscono anche che la foresta pluviale si sia ritirata in risposta al raffreddamento dell'età glaciale, per espandersi nuovamente nei periodi interglaciali più caldi, come quello che stiamo attraversando. In definitiva, la situazione sarebbe inversa a quella ipotizzata dalla teoria dei rifugi: alcune parti dei bassopiani avrebbero servito da riserva relativamente calda e umida per gli alberi della foresta pluviale mentre i rilievi, originariamente ritenuti rifugi, sarebbero divenuti inospitali per il freddo.

Che cosa può spiegare la diversità e le divergenze delle specie riconosciute dai biogeografi? Cambiamenti climatici tra i periodi glaciali e quelli interglaciali possono essere ancora ritenuti fattori importanti, anche se non come previsto dai sostenitori della teoria dei rifugi. Ritengo che la migrazione delle specie della foresta pluviale verso zone più elevate nei periodi interglaciali caldi come quello attuale avrebbe consentito la formazione di nuove associazioni di specie. Popolazioni provenienti dai bassopiani relativamente caldi delle età glaciali avrebbero potuto adattarsi alle nuove condizioni dei rilievi su cui venivano a trovarsi e divergere evolutivamente dai loro antenati, come accadde agli uccelli e alle farfalle che oggi vivono sugli altopiani.

Un'altra parte della risposta ha certamente a che fare con le variazioni climatiche e della geografia dell'Amazzonia. Il gioco delle speciazioni si è sempre svolto su un «tavolo» immenso: il bacino del Rio delle Amazzoni è infatti vasto quasi come gli Stati Uniti continentali. Questo tavolo da gioco è a prima vista uniforme in ogni sua parte - e in realtà le varie parti hanno alcune specie in comune - ma il territorio nel suo complesso non è occupato da un unico ecosistema monolitico. Piuttosto, ci sono ampie differenze regionali e locali nella piovosità, nell'andamento delle stagioni, nel suolo e nella predisposizione alle inondazioni, tutti fattori che possono influenzare la mescolanza e l'evoluzione delle specie in una data area.

Joseph H. Connell, dell'Università della California a Santa Barbara e, indipendentemente, Stephen P. Hubbel, ora alla Princeton University, avanzarono un'altra teoria per spiegare la diversificazione fra le specie, quella delle perturbazioni moderate. Secondo tale ipotesi, la più ampia varietà di specie non dovrebbe trovarsi là dove il clima è stabile, ma dove le modificazioni ambientali sono frequenti, ma non eccessive.

L'ipotesi delle perturbazioni moderate parte dal presupposto che catastrofi di immane entità portino a estinzioni su larga scala. Ciò è quanto accadrebbe se un asteroide colpisse la Terra, come di fatto può essere avvenuto all'epoca dell'estinzione dei dinosauri. Ma catastrofi di minore entità e a carattere locale, come uragani e inondazioni, non provocano generalmente l'estinzione di intere specie. Piuttosto, eliminando parte di una specie dominante, le catastrofi minori impediscono che competizioni del tipo «a un solo vincitore» giungano a conclusione (l'estinzione di una qualche specie) e danno a organismi in partenza più deboli l'opportunità di affermarsi.

Fra gli alberi della foresta pluviale si possono facilmente osservare gli effetti delle perturbazioni moderate. Grandi squarci e comunità appartenenti a successioni diverse si trovano ovunque nell'Amazzonia e in altre foreste pluviali tropicali (si veda l'articolo *La foresta pluviale tropicale* di Paul W. Richards in «Le Scienze» n. 67, marzo 1974). Ovunque, nel manto della foresta, vi siano squarci provocati dalla caduta di grandi alberi (come l'albero della gomma e la balsa) in grado di ostacolare la penetrazione della luce solare, prosperano diverse specie di piante eliofile a vita breve e le faune a esse associate. Questo processo dà inizio a una serie di successioni che, se queste specie crescono indisturbate, culmina secoli più tardi nelle associazioni di alberi giganteschi, caratteristiche dello stadio maturo delle foreste pluviali. In effetti i grandi alberi delle foreste pluviali cadono abbastanza facilmente perché hanno radici superficiali: spesso metà della massa dell'apparato radicale di queste piante è distribuita nei primi 20 centimetri del suolo; essi sono quindi soggetti a una grande pressione quando soffiano forti venti o quando lo strato superficiale del terreno viene dilavato dalle piene o dal ruscellamento.

Una crescente massa di dati suggerisce che, almeno durante l'Olocene (ma probabilmente nel corso di tutta la sua storia) l'Amazzonia sia stata sconvolta in questa o in quella regione da tempeste, fenomeni erosivi e di altro genere. In realtà, la topografia dell'Amazzonia sembra cambiare da secolo a secolo o addirittura da decennio a decennio.

Parte di quanto si sa sulla storia dell'Olocene è stato desunto dalle carote di sedimenti che i miei colleghi e io abbiamo prelevato dal fondo dei laghi dell'Ecuador. I nostri dati più antichi provengono da laghi formatisi in crateri prodotti da esplosioni vulcaniche, gli unici laghi vulcanici conosciuti in Amazzonia. Uno di questi laghi, chiamato Kumpak[a] nella lingua Shuar della popolazione locale, ha un'acqua torbida, povera di ossigeno, come del resto avviene in tutti i laghi alimentati da corsi d'acqua nel bacino del Rio delle Amazzoni. L'altro lago, Ayauch[i], è limpido e azzurro, con acque ossigenate fino in profondità, quasi come un laghetto alpino, spettacolo raro e di selvaggia bellezza nel cuore della foresta tropicale.

Mark B. Bush, un ricercatore del mio laboratorio, ha portato a termine l'analisi pollinica di reperti risalenti a 7000 anni fa prelevati dal lago Ayauch[i] e ha trovato dati relativi alla foresta pluviale per gran parte di questo arco di tempo, anche se 4000 anni fa, molto tempo dopo il termine dell'età glaciale, si è avuta una prolungata siccità locale. (Altre prove trovate da Bush indicano che già 3000 anni fa le popolazioni locali coltivavano il mais ai margini della foresta. Si tratta della più antica testimonianza di coltivazione di mais dell'Amazzonia.)

Kam-biu Liu, attualmente alla Louisiana State University, ha ricostruito un altro tipo di storia dai sedimenti prelevati dal Kumpak[a]. Questi presentano bande di differente tessitura dalle quali si può trarre la conclusione che piogge torrenziali abbiano periodicamente dilavato le sponde riversando nuovi sedimenti nel lago. Le tempeste avrebbero

Il fiume Napo, in Ecuador, è uno dei molti corsi d'acqua che, erodendo le proprie sponde e divellendo gli alberi, modificano costantemente la topografia della foresta pluviale. Secondo l'autore l'Amazzonia è un ecosistema adattato ad aggressioni di questo tipo, che non estinguono le forme di vita, ma non adattato alla catastrofe che le viene oggi inflitta con la deforestazione a opera dell'uomo.

cioè svolto un'azione di rimodellamento del territorio attorno al lago Kumpak[a], almeno negli ultimi 5000 anni.

Sempre nella foresta pluviale dell'Ecuador, ma nella zona settentrionale, carote estratte da quattro laghi di bassopiano, abbandonati ormai da secoli dai loro immissari, testimoniano chiaramente almeno un periodo protratto di uragani eccezionalmente violenti e perturbazioni della foresta nella sua parte occidentale. I campioni di sedimento provenienti da ciascun lago erano sormontati da uno strato di *gyttja*, un caratteristico fango lacustre, che iniziò a depositarsi circa 800 anni fa, secondo una datazione al radiocarbonio. Al di sotto di questo strato in tutti e quattro i laghi si trovava tipico sedimento di fiume depositato in modo continuo a partire da circa 1300 anni fa. Tuttavia, i laghi vennero abbandonati definitivamente circa 800 anni fa dai loro immissari.

Abbiamo ipotizzato che tra 800 e 1300 anni fa piogge estremamente intense che interessarono le montagne a occidente abbiano causato notevoli inondazioni costringendo i fiumi a rioccupare i loro vecchi alvei nei bassopiani che ne risultarono ampiamente allagati. L'analisi pollinica dell'antico sedimento fluviale indica che foreste corrispondenti agli stadi iniziali della successione si insediarono su provvisori banchi di sabbia che si formavano quando il livello delle inondazioni, che presumibilmente abbattevano gli alberi maturi, temporaneamente si abbassava.

Marcia L. Absy dell'Istituto amazzonico per la pesca di Manaus, in Brașile, ha trovato prove di inondazioni (che ben corrispondono nella datazione) anche in laghi nel cuore dell'Amazzonia. La Absy ha carotato cinque laghi *varzea* (che, alimentati dai fiumi durante le stagioni più piovose, si riducono lentamente per evaporazione durante le stagioni più secche) appartenenti a differenti bacini idrografici prossimi a Manaus; di essi fanno parte laghi «relitti», abbandonati da fiumi che provenengono da nord dagli aridi altopiani della Guiana, da sud dal Mato Grosso e da ovest dalle Ande ecuadoriane e peruviane. I dati raccolti dalla Absy mostrano che solo i sedimenti dei laghi dei bacini di drenaggio occidentali recano indizi di avvenute inondazioni, mentre non ne contengono quelli situati lungo i corsi d'acqua che scorrono verso nord e verso sud. Il periodo delle tempeste sarebbe così stato un fenomeno locale della parte occidentale del bacino e avrebbe avuto effetti a lungo raggio solamente sulle regioni solcate dai ricchi tributari del Rio delle Amazzoni.

L'Amazzonia occidentale ha quindi subito nell'Olocene almeno un inconsueto e drastico cambiamento climatico, i cui effetti si sono protratti per quasi un millennio. Senza alcun dubbio nel turbolento bacino rimangono da scoprire altri eventi climatici dall'effetto a lungo termine.

Se si considerano scale temporali dell'ordine di alcuni secoli risultano particolarmente evidenti gli effetti di perturbazione causati dall'erosione a opera dei corsi d'acqua. Alcuni fiumi hanno un rapido corso e pertanto rapida è la loro azione erosiva; altri scorrono lentamente, ma nondimeno sono capaci di erosione. Molti fiumi cambiano alveo divellendo gli alberi lungo il loro percorso e lasciandosi dietro sedimenti pronti per essere colonizzati. Tutti sono soggetti a catastrofiche piene nei periodi di pioggia, che modificano il paesaggio. Da questo punto di vista i torbidi tributari occidentali dell'Amazzonia sono particolarmente attivi: in effetti, secondo diverse stime l'80 per cento dei solidi sospesi nel tratto più a valle del Rio delle Amazzoni proviene da occidente.

Sulla base di valutazioni dei sedimenti trasportati ogni anno alla foce del Rio delle Amazzoni, il mio gruppo di lavoro ha tentato di calcolare la velocità a cui i grandi fiumi che scendono dalle Ande stanno erodendo la regione occidentale. Con estrapolazioni forse un po' troppo ardite si è calcolata un'asportazione dello strato superficiale di parecchi centimetri al secolo. Forse metà dello spessore del terreno in cui penetrano le radici di un tipico albero amazzonico maturo può essere erosa in un tempo inferiore alla vita media dell'albero stesso.

Jukka S. Salo e collaboratori dell'Università di Turku, in Finlandia, lavorando nell'Amazzonia peruviana, hanno ottenuto prove migliori sui danni causati dall'erosione nella parte occidentale del bacino. Con l'aiuto di immagini da satellite hanno realizzato mappe di tipi differenti di foresta, comprendenti sia comunità pioniere di una successione (che tipicamente crescono su sedimenti che i fiumi depositano nel corso dello spostamento dei propri alvei) sia comunità degli stadi maturi (che diventano stabili solo secoli dopo che un alveo di fiume è stato abbandonato). Sulla base di queste mappe i ricercatori hanno stabilito che negli ultimi due secoli un quarto della foresta peruviana è stato distrutto e ricostituito con nuove comunità.

I fiumi modificano il terreno anche su scale temporali minori. Con fotografie aeree riprese a distanza di 13 anni, Salo e collaboratori hanno dimostrato che un solo piccolo fiume aveva rimaneggiato il 3,7 per cento del suo bacino in questi pochi anni; in media aveva eroso 12 metri di terreno per anno.

La maggior parte della gente non si stupisce nell'apprendere che la foresta amazzonica subisce azioni di disturbo da parte di tempeste, alluvioni ed erosione, ma pochi si aspettano che una foresta pluviale possa essere sconvolta da incendi naturali. Recentemente Robert L. Sanford dell'Università della California a Berkeley e altri ricercatori hanno reperito strati di carbone di legna in cave di terra nel Venezuela meridionale, proprio all'estremità settentrionale del bacino amazzonico. La datazione con il radiocarbonio ha mostrato che alcuni campioni risalivano a 6000 anni fa, a prima, cioè, della presunta comparsa di esseri umani nella regione. Almeno alcuni degli strati di carbone devono dunque essere attribuiti a incendi spontanei.

Che cosa provocherebbe questi incendi? Si possono considerare gli alberi della foresta pluviale come trappole energetiche raffreddate ad acqua. Nel corso della giornata gli alberi assorbono un'intensa radiazione solare sulla vasta superficie delle loro foglie ampie e sottili, e possono svolgere questa attività solo se hanno la possibilità di dissipare calore facendo evaporare immense quantità di acqua. Sanford ritiene possibile che in un unico mese senza pioggia gli alberi potrebbero disperdere tutta l'acqua disponibile per le loro radici superficiali; conseguentemente le foglie si surriscalderebbero e appassirebbero. Il Sole potrebbe allora inaridire la lettiera ai piedi degli alberi privi di foglie, rendendo possibile la combustione spontanea o provocata da fulmini. La casualità delle condizioni atmosferiche potrebbe produrre nella foresta pluviale brevi periodi di siccità come questi una o due volte nell'arco di diverse migliaia di anni.

Risulta evidente che il bacino dell'Amazzonia è stato in realtà sempre soggetto ad azioni di disturbo. Se il nostro insieme di dati è rappresentativo della regione, nelle età glaciali il territorio si è raffreddato e la foresta si è ritirata; in epoche più calde la foresta amazzonica ha probabilmente subito perturbazioni del tipo identificato nei sedimenti lacustri e in altri reperti di età olocenica. Zone diverse in tempi diversi sono investite da tempeste, inondazioni ed erosione o da incendi nei rari periodi in cui non si hanno piogge per molti giorni di seguito. Probabilmente ben poche zone di foresta riescono a rimanere indisturbate per più di un secolo o due.

Che cosa ci suggerisce tutto ciò riguardo al futuro dell'ecosistema amazzonico, ora fortemente minacciato dall'azione umana? L'accumularsi di specie in un luogo soggetto a continui cambiamenti indica che fauna e flora sono in grado di sopportare una certa attività umana, se questa è paragonabile alle perturbazioni moderate che garantiscono la sopravvivenza di un adeguato numero di forme di vita. Va tuttavia chiaramente detto che nessun evento della storia dell'Amazzonia è paragonabile ai drastici stravolgimenti che vengono oggi inflitti dall'azione dell'uomo all'intero sistema. Questa è paragonabile solo alle catastrofi naturali che nel passato hanno portato all'estinzione di un gran numero di specie.

Oltre a ciò, le specie animali di maggiori dimensioni che abitano nella foresta pluviale non possono sopravvivere alla caccia con le moderne armi da fuoco. In effetti, i primati erbivori e i bradipi che si nutrono nella zona «di baldacchino», nonché i volatili loro predatori come l'arpia (una macchina da predazione così potente da strappare le scimmie dal loro habitat), sono disperatamente indifesi nei confronti dei fucili da caccia. Un singolo uomo armato di un fucile calibro 16 può uccidere in un anno tutte le arpie e tutti i primati meno mobili che vivono in un'area di 10 chilometri, e questo è ciò che hanno fatto migliaia di cacciatori. La sola speranza per questi animali è rappresentata dalla creazione di aree protette; i governi dei diversi stati amazzonici stanno organizzando riserve faunistiche, mentre sono in corso ricerche per stabilire le dimensioni più adeguate per questi rifugi.

Per quanto riguarda gli alberi la storia è diversa: se continuerà l'attuale tendenza, gran parte del bacino amazzonico sarà trasformata in pascolo a mano a mano che il territorio verrà deforestato per allevare bestiame. La sola speranza per la flora e per gli insetti risiede probabilmente nello sviluppo di forme di sfruttamento meno dissennate, che non infliggano a quanto rimane della foresta pluviale danni più gravi di quelli a cui essa si è adattata. Forse parti di foresta potrebbero essere utilizzate per un turismo di élite e case di riposo o per installarvi piccole attività artigianali o industriali che non richiedano enormi quantità di energia e non inquinino l'ecosistema. La storia suggerisce che parti dell'Amazzonia possono essere sfruttate produttivamente senza causare estinzioni di massa, purché le esigenze dell'ecosistema vengano anteposte a qualunque iniziativa.

BIBLIOGRAFIA

DAMUTH JOHN E. e FAIRBRIDGE RHODES W., *Equatorial Atlantic Deep--Sea Arkosic Sands and Ice-Age Aridity in Tropical South America* in «Geological Society of America Bulletin», 81, n. 1, gennaio 1970.

IRION G., *Sedimentation and Sediments of Amazonian Rivers and Evolution of the Amazonian Landscape since Pliocene Times* in *The Amazon: Limnology and Landscape Ecology of a Mighty Tropical River and Its Basin* a cura di Harald Sioli, Dr. W. Junk Publishers, 1984.

WHITMORE T. C. e PRANCE G. T. (a cura), *Biogeography and Quaternary History in Tropical America*, Oxford University Press, 1987.

COLINVAUX PAUL, *Amazon Diversity in Light of the Paleoecological Record* in «Quaternary Science Reviews», 6, n. 2, 1987.

SALO JUKKA, *Pleistocene Forest Refuges in the Amazon: Evaluation of the Biostratigraphical, Lithostratigraphical and Geomorphological Data* in «Annales Zoologici Fennici», 24, pp. 203-211, 1987.

Le raganelle marsupiali

Queste specie tropicali in cui l'incubazione delle uova avviene sul dorso della femmina, spesso in un'apposita «tasca», presentano adattamenti riproduttivi singolari, che ricordano quelli dei mammiferi e degli uccelli

di Eugenia M. del Pino

I mammiferi mettono al mondo piccoli vivi e li allattano; gli uccelli depongono uova e le covano. Gli anuri, come tutti gli anfibi, depositano le uova nell'acqua e da esse nascono larve acquatiche, i girini. Come indicano queste generalizzazioni, le modalità di riproduzione hanno un'importanza che trascende la maggior parte delle differenze di conformazione fisica all'interno di ogni classe animale: per questo motivo i pipistrelli sono classificati assieme agli elefanti, i colibrì assieme agli struzzi.

Nelle foreste pluviali e sulle montagne dell'America Centrale e delle zone tropicali dell'America Meridionale vivono circa 60 specie di raganelle arboricole che sembrano smentire questa regola. Da adulte, esse rassomigliano alle normali raganelle, ma il loro meccanismo riproduttivo non è affatto simile a quello degli altri anfibi. Le uova infatti non si sviluppano nell'acqua ma sul dorso della madre, spesso all'interno di una speciale tasca, una sorta di marsupio. In molte specie i piccoli ne escono come ranocchietti già pienamente formati, in altre come girini o a uno stadio più avanzato del loro sviluppo.

Queste particolari modalità riproduttive hanno reso necessari alcuni adattamenti specializzati. In posizione protetta sul dorso della madre, l'embrione può svilupparsi a una velocità relativamente bassa. A sua volta, questo sviluppo prolungato nel tempo implica che l'embrione debba essere rifornito di sostanze nutritive e sia inoltre in grado di scambiare gas e liquidi con l'ambiente che lo circonda. Tutto ciò ha imposto cambiamenti nella fisiologia della femmina, nella struttura e nelle caratteristiche molecolari dell'uovo e nelle modalità di sviluppo embrionale. Molti di questi adattamenti specializzati permettono confronti tra le raganelle che incubano le uova e animali da esse molto distanti dal punto di vista evolutivo.

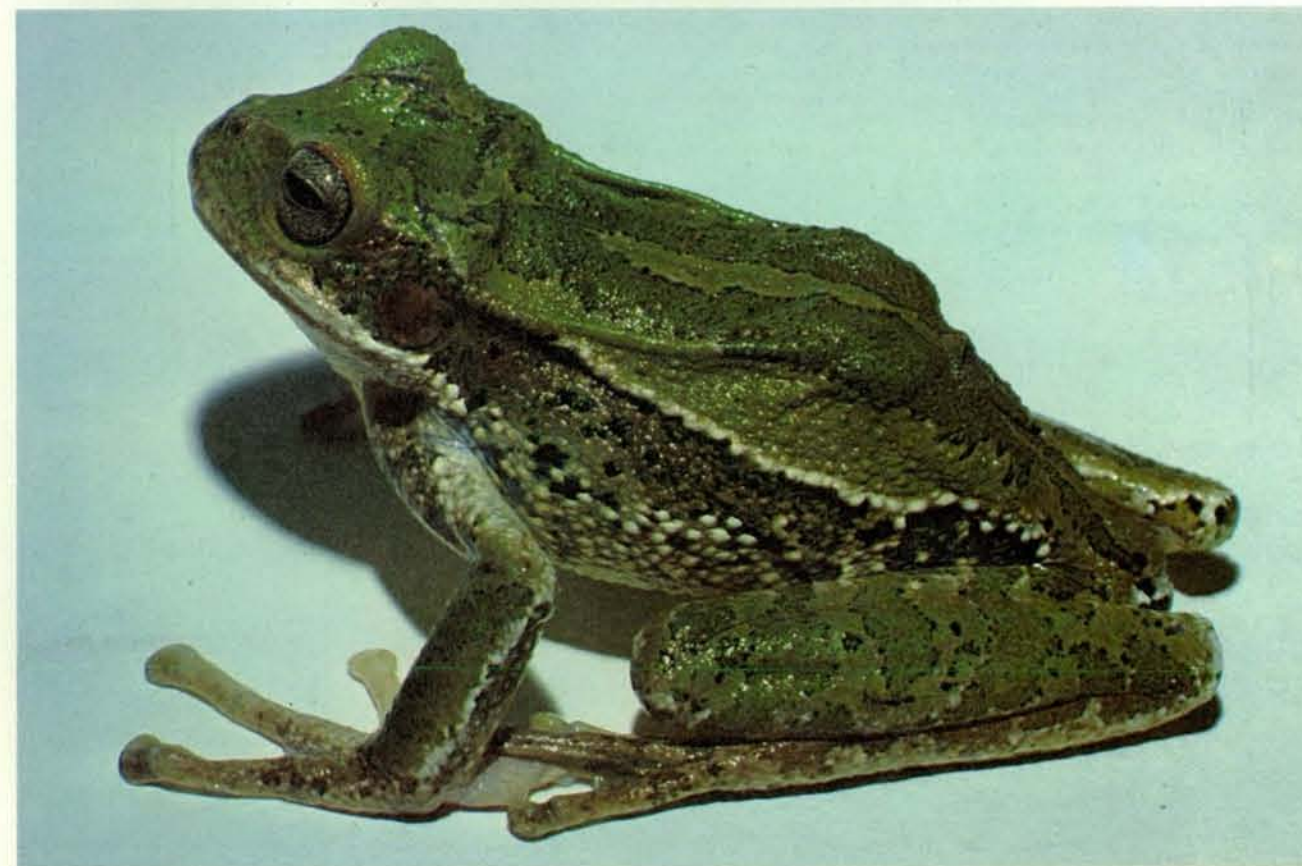
La tasca di incubazione di queste raganelle ricorda naturalmente quella dei marsupiali. La struttura interna della tasca e gli ormoni che regolano l'incubazione nelle raganelle marsupiali mostrano parallelismi con gli apparati e le secrezioni ormonali che presiedono alla gestazione nei mammiferi. L'uovo, che è grande e ricco di tuorlo, assomiglia di più alle uova degli uccelli o dei rettili che non a quelle degli altri anfibi e la stessa cosa, sotto alcuni aspetti, si può dire per lo sviluppo embrionale. Tuttavia, se consideriamo altre caratteristiche riproduttive, le raganelle marsupiali presentano notevoli singolarità: questo fatto ci ricorda che la somiglianza con altre classi di animali non è il risultato di una discendenza diretta lungo un particolare ramo evolutivo, ma di un fenomeno di convergenza che si verifica tra rami separati da centinaia di milioni di anni di storia evolutiva.

Nel 1972, quando iniziai la mia attività di insegnamento presso la Pontificia Università cattolica dell'Ecuador a Quito, le particolari caratteristiche riproduttive delle raganelle che incubano le uova erano conosciute solo in modo assai approssimativo. Per me questi anfibi non rappresentavano niente di più che una curiosità da libro di testo. Poco dopo essermi trasferita a Quito, però, scoprii che una specie di raganella marsupiale, *Gastrotheca riobambae*, viveva nei giardini dell'università. Cominciai pertanto a interessarmi a questa raganella e a studiarne le modalità di riproduzione, e mi misi alla ricerca di esemplari di altre specie affini.

G. riobambae è una fra le poche specie di raganelle che incubano le uova che siano riuscite a insediarsi (un evento raro tra gli anuri in generale) nell'ambiente d'alta montagna di Quito, città situata a quasi 3000 metri sul livello del mare. La maggior parte di queste raganelle arboricole vive nel baldacchino delle foreste pluviali dei bassopiani ecuadoriani e in altre regioni dell'America Meridionale e Centrale. Nello studio di questa specie mi sono basata sia su esemplari conservati in museo sia su animali catturati vivi e portati nel mio laboratorio da William E. Duellman, Linda Trueb, John Simmons e altri ricercatori del Museum of Natural History dell'Università del Kansas.

La foresta pluviale tropicale, ambiente nel quale le raganelle che incubano le uova presentano la maggior varietà di specie, è probabilmente anche il luogo in cui è avvenuta la loro evoluzione. È facile intuire come le pressioni selettive di questo ambiente abbiano potuto favorire l'evoluzione di questa modalità particolare di incubazione delle uova. Il gran numero di specie di anuri presenti nella foresta tropicale fa sì che i siti più favorevoli per la riproduzione siano oggetto di competizione intensa. (In appena tre chilometri quadrati di foresta presso Santa Cecilia, in Ecuador, Duellman e la sua assistente Martha L. Crump hanno trovato ben 81 specie di anuri, circa lo stesso numero delle specie che vivono in tutta l'America Settentrionale!) Inoltre molti predatori minacciano le uova e le larve degli anuri, che si trovano senza alcuna difesa negli stagni e nei corsi d'acqua.

Di conseguenza, molte specie di anuri tropicali hanno adottato nuovi siti per la deposizione delle uova. Alcune le depositano su foglie al di sopra di uno stagno o di un corso d'acqua; alla schiusa la pioggia trascina nell'acqua i girini. Altri anuri nascondono le uova, da cui nascono ranocchietti già sviluppati, sotto i detriti che ricoprono il terreno della foresta. Ma il più complicato meccanismo di incubazione delle uova sul dorso della femmina si è evoluto in una sottofamiglia di ilidi (raganelle arboricole), quella degli emifrattini.

Si tratta di uno sviluppo evolutivo molto antico. Brian E. Scalan e Linda R. Maxson dell'Università dell'Illinois a Urbana-Champaign, in collaborazione con Duellman, hanno ricostruito nel

Una tasca di incubazione (marsupio) piena di embrioni gonfia a dismisura questo esemplare di *Gastrotheca riobambae*, una raganella marsupiale dell'Ecuador (*in alto*). Dopo l'uscita di 218 girini attraverso l'apertura della tasca, nella parte bassa del dorso, la rana riassume dimensioni normali (*in basso*). Le fotografie sono di Charles W. Myers dell'American Museum of Natural History.

tempo l'evoluzione delle raganelle che incubano le uova mediante il confronto tra le sequenze amminoacidiche di certe proteine in specie diverse. Supponendo che le sequenze di amminoacidi siano variate a una velocità costante nel corso dell'evoluzione, la differenza tra due sequenze qualsiasi dovrebbe essere proporzionale alla distanza evolutiva tra due organismi e cioè al tempo che li separa da un antenato comune. I confronti tra le proteine indicano che i rami evolutivi cominciarono a separarsi tra 40 e 80 milioni d'anni fa; presumibilmente l'incubazione delle uova era già presente nel ceppo ancestrale.

Attualmente esiste una notevole varietà nei sette generi - che comprendono circa 60 specie - di raganelle che incubano le uova. Tre generi (*Stefania*, *Cryptobatrachus* ed *Hemiphractus*) non possiedono marsupio; le uova aderiscono al dorso della madre, dove si sviluppano e alla schiusa danno vita a ranocchietti già formati. Gli altri quattro generi, dei quali mi occuperò in particolare in questo articolo (*Fritziana*, *Flectonotus*, *Gastrotheca* e *Amphignathodon*), sono vere raganelle marsupiali, fornite di tasche d'incubazione.

Le variazioni nella forma della tasca nei diversi generi fanno pensare che questa struttura possa essere derivata da pieghe longitudinali presenti sulla pelle del dorso. In alcune specie di *Fritziana*, per esempio, gli embrioni in via di sviluppo sono trattenuti dalle pieghe ma per il resto sono scoperti. In altre specie di *Fritziana* e nel genere *Flectonotus* le pieghe diventano falde che s'incontrano sulla linea mediana del dorso, racchiudendo gli embrioni in una tasca costituita da due compartimenti laterali. Il marsupio si trasforma in una struttura più elaborata in *Amphignathodon* e in molte specie di *Gastrotheca*: ha ancora due camere laterali, ma l'unica apertura è un pertugio nella parte bassa del dorso, appena al di sopra della cloaca (l'orifizio comune all'apparato riproduttore ed escretore). Le specie d'alta montagna di *Gastrotheca* (tra cui *G. riobambae*) possiedono quella che può essere definita la forma più evoluta di tasca incubatrice, in cui i due compartimenti laterali sono fusi in uno solo.

L'evoluzione del marsupio ha portato a nuovi comportamenti riproduttivi. Gli anuri si accoppiano nel momento in cui la femmina sta deponendo le uova, che vengono fecondate a mano a mano che escono dalla cloaca; di solito l'accoppiamento ha luogo nell'acqua e qui vengono lasciate le uova. Le raganelle marsupiali invece si accoppiano sulla terraferma, ma devono affrontare un altro problema: come far arrivare le uova dalla cloaca al marsupio.

Nelle tre o quattro specie di cui è stato osservato il comportamento sessuale, è il maschio a eseguire attivamente il trasferimento. Durante l'accoppiamento di *G. riobambae*, per esempio, il maschio si arrampica sul dorso della femmina e apre il marsupio con le zampe posteriori. Appena le uova emergono dalla cloaca della femmina, una o due al minuto, il maschio le afferra una per volta con la zampa e le sospinge nel marsupio: l'uovo viene fecondato in questo tragitto dallo sperma che il maschio ha depositato sul dorso della femmina. Questa operazione continua finché in media 130 uova sono state stipate nel marsupio. Successivamente la femmina preme il dorso contro una pietra per assestare le uova così «insaccate» in numerosi strati uniformi.

In molte specie di raganelle marsupiali passano parecchi mesi prima che avvenga la schiusa delle uova. Quando è il momento di uscire dal marsupio, i giovani anfibi - girini più o meno sviluppati in *Fritziana*, *Flectonotus* e in alcune specie di *Gastrotheca*, e ranocchietti in altri casi - hanno ancora bisogno d'aiuto da parte della madre. Una femmina di *G. riobambae*, per esempio, lascia i girini neonati in pozze o stagni aggrappandosi a una pietra con le zampe anteriori, mentre piega le posteriori per raggiungere l'apertura della tasca. In questo modo estrae i girini uno alla volta, impiegando due o più giorni per mettere al mondo l'intera nidiata.

Nelle specie che danno vita a ranocchietti già completamente sviluppati, il processo non deve svolgersi necessariamente nei pressi dell'acqua. Mentre sono nella tasca, gli embrioni di queste raganelle sviluppano una versione rudimentale di molte caratteristiche che si osservano normalmente nei girini. Al momento della nascita, la metamorfosi risulta completa: questi anfibi quindi «saltano» virtualmente la fase acquatica. Presumibilmente molte di queste raganelle riescono a compiere l'intero ciclo vitale senza mai scendere dal baldacchino della foresta.

Gli emifrattini, o raganelle che incubano le uova, costituiscono una sottofamiglia delle comuni raganelle, appartenenti alla famiglia degli ilidi. Basandosi su confronti fra le proteine e le caratteristiche riproduttive di queste raganelle, William E. Duellman del Museum of Natural History dell'Università del Kansas e Richard J. Wassersug della Dalhousie University nella Nuova Scozia hanno proposto lo schema di rapporti evolutivi rappresentato nell'illustrazione. Tre dei sette generi non possiedono marsupio: gli embrioni in via di sviluppo aderiscono semplicemente al dorso della madre. I diversi generi vengono distinti anche in base al fatto che i piccoli nascano come ranocchietti già completamente formati oppure come girini a uno stadio avanzato di sviluppo (*disegni più piccoli*). In un genere - *Gastrotheca* - la forma alla nascita varia da specie a specie. Le quattro raganelle non sono rappresentate in scala uguale.

L'incubazione delle uova da parte delle raganelle marsupiali è accompagnata da un singolare comportamento riproduttivo. Durante la copula, nella specie *Gastrotheca riobambae*, la femmina depone le uova mentre il maschio la afferra da dietro e inserisce le dita delle zampe posteriori nel marsupio (*a sinistra*). Via via che le uova emergono dalla cloaca sollevata, il maschio le spinge nel marsupio; le uova vengono fecondate dallo sperma deposto in precedenza sul dorso della femmina. Nel giro di alcune ore vengono trasferite 100 o più uova. Lo sviluppo embrionale dura più di tre mesi, durante i quali la femmina diviene così gonfia che i suoi movimenti risultano impacciati (*al centro*). Essa mette al mondo i girini nell'acqua, facilitandone l'uscita con le zampe posteriori (*a destra*).

L'ambiente protetto del marsupio fa sì che lo sviluppo embrionale possa procedere molto più lentamente nelle raganelle marsupiali rispetto alle specie in cui le uova sono esposte e vulnerabili. L'incubazione in *G. riobambae*, per esempio, dura tra i 100 e i 120 giorni, mentre occorre solo un giorno e mezzo per la schiusa delle uova, deposte nell'acqua, di *Xenopus laevis* (una specie di anfibio fornito di artigli che vive in Africa), che permette un confronto adeguato poiché è stato studiato in modo completo. Come viene tenuto in vita l'uovo nel corso della lunga incubazione? Un embrione di anuro che si sviluppa in pochi giorni dando vita a un girino in grado di condurre un'esistenza autonoma non necessita di speciali meccanismi per lo scambio di gas, di liquidi e di altre sostanze con l'ambiente circostante. Nelle raganelle marsupiali, invece, ci si dovrebbe aspettare che tanto la madre quanto l'embrione siano adattati a questi scambi.

Nell'intervallo tra due stagioni riproduttive il marsupio vuoto presenta una parete interna di aspetto simile alla normale pelle della raganella. Tuttavia, durante l'incubazione, su questa parete si forma un rivestimento sottile e riccamente irrorato di vasi sanguigni. Questo rivestimento interno, che viene eliminato successivamente alla schiusa delle uova, si adatta alla forma di queste ultime, e va a costituire una camera singola per ciascun uovo. All'interno della camera d'incubazione l'embrione è circondato da una membrana vascolarizzata propria: un complesso specializzato di branchie dette «branchie a campana».

Le branchie a campana, sconosciute in altre specie di anuri, iniziano a svilupparsi dopo parecchie settimane d'incubazione. Inizialmente si presentano come semplici pellicole di tessuto che si estendono su entrambi i lati della testa dell'embrione; in seguito, in alcune specie, esse si fondono e avvolgono completamente l'embrione, formando una sacca piena di liquido in cui l'embrione rimane immerso. Alla nascita, le branchie a campana vengono riassorbite e quindi eliminate.

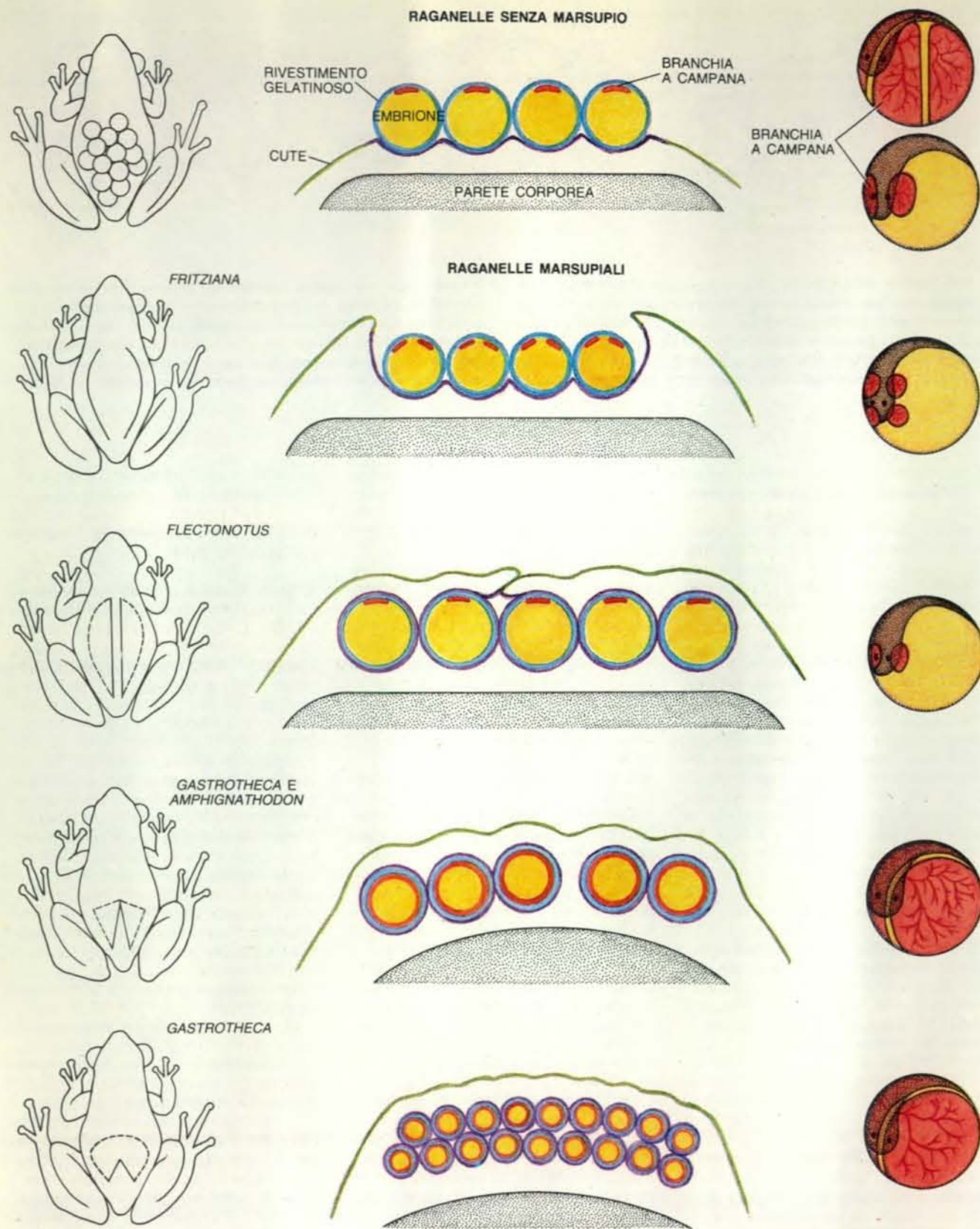
La madre e gli embrioni contenuti nella tasca si trovano quindi in stretto contatto fisiologico. Solo la capsula gelatinosa dell'uovo separa il rivestimento interno della camera ovulare dalle branchie a campana. Nelle raganelle marsupiali lo strato gelatinoso è eccezionalmente sottile: mentre negli altri anuri esso raggiunge lo spessore di alcuni millimetri, nelle raganelle marsupiali può essere anche mille volte più sottile, fino allo spessore di appena qualche micrometro.

Sotto molti punti di vista questo apparato rassomiglia a quello deputato alla gestazione nei mammiferi. La camera piena di liquido in cui avviene lo sviluppo dell'embrione ricorda il sacco di liquido amniotico che circonda gli embrioni dei mammiferi. Inoltre, cosa ancora più importante, lo stretto contatto tra il circolo sanguigno della madre e quello dell'embrione, attraverso le membrane giustapposte, fa pensare al ruolo svolto nei mammiferi dalla placenta, l'organo che collega l'apparato circolatorio della madre e quello dell'embrione durante la gravidanza.

La placenta permette alla madre di scambiare gas, liquidi, sostanze nutritive e rifiuti organici con l'embrione. È presumibile che anche nelle raganelle marsupiali si stabilisca un passaggio di questo genere attraverso le due membrane. Bertha Escobar e io abbiamo pesato embrioni prelevati da femmine di *G. riobambae* a vari stadi d'incubazione per verificare quali sostanze essi riescano ad assorbire dalla madre. Il peso umido degli embrioni è aumentato di tre volte nel corso dell'incubazione, ma il peso secco è rimasto relativamente costante. Ne abbiamo dedotto che gli embrioni hanno assorbito acqua e forse hanno scambiato gas con la madre; presumibilmente le sostanze nutritive che hanno favorito il loro sviluppo provenivano invece in gran parte dal tuorlo, la sostanza di riserva dell'embrione stesso.

Quali stimoli ormonali preparano la tasca per l'incubazione? Nel tentativo di dare una risposta a questa domanda, le mie colleghe Carmen de Albuja, Margarita Campos e io ci siamo imbattute in ulteriori analogie tra l'incubazione delle uova nelle raganelle e la gestazione nei mammiferi. In questi ultimi, diversi ormoni entrano in gioco nel predisporre l'utero alla gravidanza. Le gonadotropine vengono secrete dall'ipofisi e stimolano il follicolo ovarico, l'aggregato di cellule che circonda l'oocita in via di sviluppo (ossia la cellula che, una volta maturata, costituirà l'uovo fecondabile). In risposta a questo stimolo, il follicolo libera ormoni propri - estrogeni e progesterone - che inducono una proliferazione della mucosa uterina in modo tale da predisporla a ricevere la cellula uovo fecondata.

Nella maggior parte degli anuri esiste un meccanismo ormonale simile, a due stadi, ma il suo effetto principale è quello di promuovere la crescita e la maturazione degli oociti. Le gonadotropine stimolano i follicoli ovarici a liberare estrogeni e progesterone; il progesterone induce la serie di cambiamenti cellulari che trasformano un oocita in un uovo. Nelle raganelle che incubano le uova, tuttavia, questo meccanismo ormonale ha assunto un'ulteriore funzione, quella di preparare il marsupio per l'incubazione.

Quando la stagione della riproduzione si approssima, in *G. riobambae* e in altre specie di raganelle marsupiali si può osservare un mutamento preliminare indispensabile. Per la maggior parte dell'anno l'apertura del marsupio è dilatata, ma quando la femmina si appresta alla riproduzione tale apertura si restringe. Il mar-

Nelle raganelle che incubano le uova si stabilisce tra la madre e gli embrioni uno stretto contatto. Per ciascun genere di raganelle è raffigurata la forma del marsupio oppure l'area del dorso alla quale aderiscono le uova (*a sinistra*); nelle sezioni verticali (*al centro*) viene evidenziato il rapporto tra la madre e gli embrioni. Una membrana fortemente vascolarizzata (*in viola*) avvolge ciascun uovo e solo il sottile strato di rivestimento gelatinoso dell'uovo (*in azzurro*) la separa dalle membrane vascolarizzate proprie dell'embrione, le branchie a campana (*in rosso*). La forma della branchie varia: in alcuni generi esse coprono completamente l'embrione (*a destra*). Nelle raganelle marsupiali la comunicazione tra i tessuti della madre e le branchie a campana dell'embrione permette loro di scambiare gas e liquidi: è proprio lo stesso tipo di comunicazione esistente tra madre e figlio nei mammiferi, tramite la placenta.

supio rimane chiuso per tutto il periodo dell'incubazione, in modo da trattenere all'interno gli embrioni, fino al momento in cui i girini o i ranocchietti sono pronti per venire alla luce.

Il controllo ormonale della chiusura del marsupio può essere studiato nei particolari inoculando gli ormoni uno per volta e osservandone gli effetti separati. Se a una raganella viene praticata un'iniezione di gonadotropina, il marsupio si chiude solo nel caso in cui le ovaie contengano grossi follicoli. Al contrario, una sola iniezione di progesterone provoca il restringimento dell'apertura del marsupio entro 24 ore, indipendentemente dalle dimensioni dei follicoli ovarici. Evidentemente l'ormone direttamente responsabile della chiusura del marsupio è il progesterone secreto dai follicoli, mentre le gonadotropine hanno l'unica funzione di stimolare la liberazione di progesterone.

Anche lo sviluppo delle camere ovulari, che avvolgono ciascun embrione in una membrana vascolarizzata, sembra essere promosso dal progesterone. Se si introducono sferette di plastica dello stesso diametro degli embrioni nel marsupio d'una femmina quando questa non è in fase d'incubazione, le sferette vengono generalmente espulse nel giro di poche ore. Tuttavia, se alla raganella viene inoculato progesterone contemporaneamente all'inserimento delle sferette nel marsupio, questo si rinchiude sopra le sferette, che vengono avviluppate dalle membrane vascolarizzate delle camere ovulari. I falsi embrioni vengono trattenuti nella tasca anche fino a una settimana.

Al momento dell'incubazione, i follicoli si sono già aperti e hanno liberato le uova mature. Allora, da dove viene il progesterone che promuove la crescita delle camere ovulari? Durante i primi stadi della gravidanza nei mammiferi, il follicolo vuoto continua a produrre progesterone anche successivamente alla liberazione della cellula uovo, stimolando in questo modo lo sviluppo della mucosa uterina. Nella maggior parte degli anuri, i follicoli rimangono nelle ovaie dopo l'ovulazione per meno di una settimana e si pensa che non abbiano una funzione ormonale. Invece in *G. riobambae* i follicoli vuoti rimangono nelle ovaie per più di un mese, arco di tempo corrispondente a un terzo dell'intero periodo di incubazione.

Per verificare se i follicoli dopo l'ovulazione abbiano effettivamente la capacità di produrre progesterone, Grace Sánchez e io abbiamo asportato le ovaie (e con esse i follicoli ancora presenti) da raganelle che stavano incubando le uova. L'ovariectomia durante le primissime settimane di incubazione provocava aborto: gli embrioni parzialmente sviluppati venivano espulsi dal marsupio dopo circa 15-30 giorni. L'asportazione delle ovaie in uno stadio più tardivo dell'incubazione non aveva invece alcuna conseguenza. Quindi, almeno per quanto riguarda i primi stadi, l'incubazione nelle raganelle marsupiali sembra avere basi ormonali simili a quelle della gravidanza nei mammiferi.

È possibile che le analogie tra mammiferi e raganelle marsupiali si estendano anche a fasi più tardive dell'incubazione. Il sostegno ormonale per gli stadi tardivi della gravidanza nei mammiferi proviene dalla placenta stessa, che secerne gonadotropine, progesterone e altri ormoni. Forse le branchie a campana e le camere ovulari nelle raganelle marsupiali hanno un ruolo analogo a quello della placenta nei mammiferi e liberano ormoni che favoriscono la fase avanzata dell'incubazione.

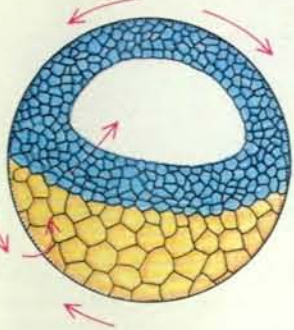
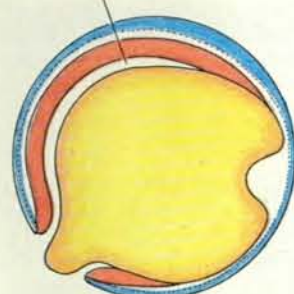
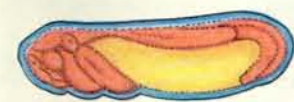
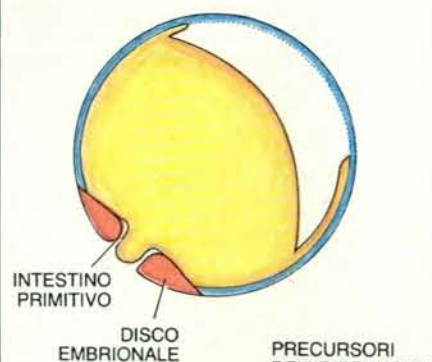
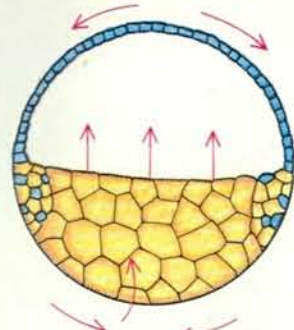
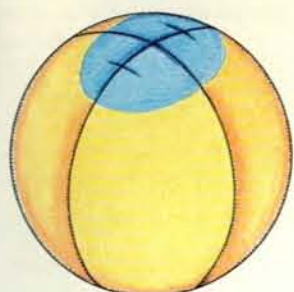
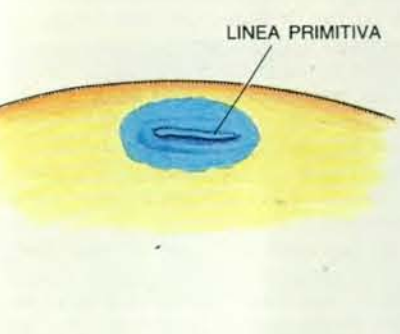
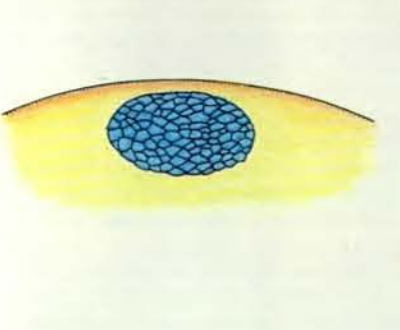
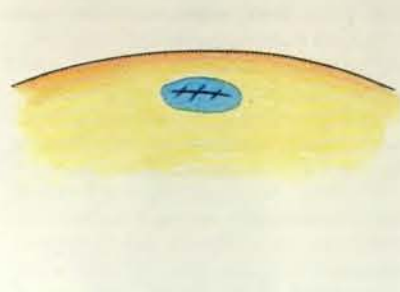
IPOFISI
GONADOTROPINE
OOCITA
FOLLICOLO
UOVO
PROGESTERONE
FOLLICOLO DOPO L'OVULAZIONE

Secondo le ricerche compiute dall'autrice e colleghi, una serie di ormoni agisce di concerto per regolare la riproduzione nelle raganelle marsupiali. Si ritiene che le gonadotropine liberate dall'ipofisi, una ghiandola situata alla base del cervello, agiscano nell'ovaio stimolando il follicolo, costituito da cellule che circondano l'oocita (il precursore dell'uovo). In risposta, il follicolo secerne progesterone, che stimola la maturazione dell'oocita e provoca anche il restringimento dell'apertura del marsupio. Dopo che l'uovo è stato liberato dall'ovaio, il follicolo vuoto può continuare a produrre progesterone, che promuove lo sviluppo delle camere vascolari che circondano gli embrioni nel marsupio. L'utero dei mammiferi viene preparato alla gestazione da un meccanismo ormonale del tutto simile.

Finora ho solo schematizzato gli adattamenti che permettono a una raganella di covare le uova all'interno del proprio corpo. Ma che significato ha per le uova delle raganelle marsupiali questa incubazione prolungata nel corpo della madre?

L'incubazione interna, che permette uno sviluppo embrionale più lento, ha una conseguenza evidente in questi anfibi: le uova hanno dimensioni eccezionalmente grandi e sono ben fornite di tuorlo, il che le rende più simili a quelle degli uccelli e dei rettili che a quelle di altri anfibi. Le uova di una specie di raganella marsupiale, *Gastrotheca cornuta*, hanno diametro di un centimetro; nelle specie affini sono frequenti uova con diametro di mezzo centimetro, mentre per esempio quelle di *Xenopus* hanno un diametro di poco superiore a un millimetro. Le riserve di tuorlo, grassi e carboidrati che determinano le dimensioni eccezionali di cui si è detto danno sostentamento all'embrione durante il suo lungo soggiorno all'interno del marsupio.

Una seconda differenza sostanziale tra le normali uova degli anuri e quelle che si sviluppano nella tasca incubatrice delle raganelle marsupiali è a livello dei meccanismi molecolari dell'oocita, il precursore dell'uovo maturo. Il rapido sviluppo che caratterizza la maggior parte degli embrioni di anuro richiede una veloce sintesi delle proteine. Una caratteristica molecolare di gran parte degli oociti di anuro permette loro di accelerare la sintesi proteica: la presenza di grandi quantitativi di RNA ribosomale (r-RNA) insieme ad altri tipi di molecole di RNA.

L'RNA ribosomale è un componente strutturale fondamentale dei ribosomi, gli organelli cellulari in cui viene operata la sintesi proteica. Dato che negli oociti della maggior parte degli anuri è

Le prime fasi dello sviluppo embrionale di *Gastrotheca riobambae* ricordano quelle degli uccelli. Gli schemi, che mettono a confronto gli stadi di sviluppo di *Gastrotheca* (*al centro*) con quelli di *Xenopus laevis* (*a sinistra*), un anuro che si riproduce nel modo usuale, e di pollo (*a destra*), sono disegnati a scale diverse: l'embrione di *Xenopus* è grande un terzo di quello di *Gastrotheca*. In *Xenopus* dopo le prime divisioni cellulari (*1*) si forma una sfera cava di cellule, la blastula (*2*), che si ripiega poi su se stessa nella gastrulazione: grandi cellule ricche di tuorlo (*in giallo*), situate su un lato dell'embrione, invadono l'interno, mentre le cellule povere di tuorlo (*in azzurro*) all'altra estremità migrano, coprendo la superficie dell'embrione. Si forma così una nuova cavità, l'intestino primitivo (*3*). L'intero embrione partecipa poi alla formazione dell'organismo (*4*). Nell'uovo di pollo, grande e pieno di sostanza vitellina, solo un piccolo disco di cellule sulla superficie del tuorlo subisce la gastrulazione e dà vita al pulcino. In *Gastrotheca* le prime fasi dello sviluppo si differenziano da quelle tipiche degli altri anuri. Come in *Xenopus*, l'intero embrione subisce la gastrulazione. Tuttavia, forse per le grandi dimensioni dell'embrione e l'alto contenuto di tuorlo, le cellule che partecipano al successivo sviluppo si concentrano su un lato, in un disco che circonda un piccolo intestino primitivo. Come nel pollo, il corpo prende forma sulla superficie del tuorlo.

contenuto un notevole quantitativo di r-RNA - circa 200 000 volte superiore a quello presente nelle normali cellule dell'organismo di un anuro - le cellule dell'embrione non hanno bisogno di sintetizzarlo prima di produrre proteine: perciò l'embrione, che è alquanto vulnerabile, può svilupparsi e uscire dall'uovo rapidamente. I mammiferi invece, il cui embrione si sviluppa protetto dal corpo della madre e può svilupparsi lentamente, immagazzinano poco r-RNA nei loro ociti.

Gli alti livelli di r-RNA riflettono a loro volta processi che avvengono nel nucleo dell'oocita di anuro. L'RNA ribosomale, come gli altri tipi di RNA nelle cellule, è trascritto dal DNA dei geni nel nucleo. Gli oociti di anuro possono accumulare quantità elevate di r-RNA perché i loro geni ribosomali sono stati amplificati, ossia duplicati molte volte. L'oocita di *Xenopus*, per esempio, contiene circa due milioni di copie di DNA ribosomale, o r-DNA. Questi geni non risiedono sui cromosomi assieme al resto dei geni della cellula, ma sono segregati all'interno del nucleo in compartimenti detti nucleoli: vi sono più di 1000 nucleoli nel nucleo di un oocita di *Xenopus* pienamente sviluppato.

Una misura della rapidità di sviluppo resa possibile dall'enorme quantità di r-RNA accumulato nell'oocita è data dal tempo nel quale avviene la gastrulazione. La gastrulazione è lo stadio dello sviluppo embrionale nel quale l'embrione ancora indifferenziato, costituito da una semplice sfera di cellule, si ripiega su se stesso per dare origine a una struttura costituita da due strati epiteliali, contenente una cavità da cui prenderà forma l'intestino. In *Xenopus* la gastrulazione avviene solo 14 ore dopo la fecondazione.

In *G. riobambae*, la specie su cui abbiamo compiuto le ricerche più dettagliate, l'intervallo tra fecondazione e gastrulazione è di 14 giorni. Mi sono chiesta se questo rallentamento nello sviluppo corrispondesse a un livello inferiore di amplificazione dei geni ribosomali e a un contenuto più basso di r--RNA. I risultati di osservazioni compiute su oociti di *Gastrotheca* hanno confermato la mia ipotesi. Michael F. Trendelenburg del Centro di ricerca sul cancro di Heidelberg e io abbiamo contato meno di 300 nucleoli negli oociti di *G. riobambae*, cioè circa un quarto del numero registrato negli oociti di *Xenopus*. Questo dato dovrebbe già essere il segnale di un livello inferiore di amplificazione dei geni ribosomali e perciò di una minore produzione di r-RNA.

Abbiamo quindi fatto ricorso a metodi molecolari per misurare il livello di amplificazione dei geni ribosomali negli oociti di *Gastrotheca*. Christine Dreyer del Max-Planck-Institut für Entwicklungsbiologie (un istituto dove si studia la biologia dello sviluppo) di Tübingen e io abbiamo esposto gli oociti a una sonda costituita da r-DNA clonato estratto da *Xenopus*. (Il DNA a filamento singolo estratto da un organismo può legarsi a sequenze corrispondenti nel DNA o nell'RNA di altri organismi e può così servire a identificarle.) La sonda ha evidenziato una piccola quantità di r-DNA.

Per scoprire se il basso livello di amplificazione genica abbia effettivamente come conseguenza la produzione di una minore quantità di r-RNA, Peter Hausen di Tübingen e io abbiamo trattato sezioni di oocita con un colorante specifico per l'RNA. La colorazione è risultata debole in *Gastrotheca*, vivida in *Xenopus*. Isolando l'RNA degli oociti e quindi utilizzando una sonda di r-DNA, Herbert Steinbeisser e Ansgar Hofmann del Centro tedesco di ricerca sul cancro e io abbiamo misurato direttamente il contenuto di r-RNA. I risultati indicano che gli oociti di *Gastrotheca* contengono effettivamente una quantità molto minore di r-RNA rispetto agli oociti di *Xenopus*.

Pertanto, anche a livello molecolare, alcune caratteristiche dello sviluppo delle raganelle marsupiali non sono simili a quelle degli anfibi, ma piuttosto a quelle dei mammiferi. Tuttavia, nei miei primi studi, avevo scoperto una caratteristica degli oociti di alcune raganelle marsupiali che è virtualmente assente non solo negli altri anuri, ma anche in tutti gli altri vertebrati. Nel 1976 il compianto Scott J. Maness della Stazione biologica Rancho Grande, in Venezuela, mi segnalò l'interessantissima raganella marsupiale *Flectonotus pygmaeus*. Non più lunga di due centimetri e mezzo, questa raganella, che vive nel baldacchino della foresta pluviale, tiene in incubazione nel suo marsupio da cinque a 11 uova e mette al mondo girini deponendoli nei pozzetti formati da foglie che raccolgono l'acqua piovana.

La caratteristica più notevole di *Flectonotus* è venuta in luce quando ho potuto osservare al microscopio un oocita aperto. Invece di un solo nucleo, si sono riversate all'esterno migliaia di strutture sferiche. Esaminando queste strutture e studiandone l'attività biochimica, il mio ex professore, Asa A. Humphries, Jr., allora alla Emory University, e io le abbiamo identificate con sicurezza come nuclei attivi dell'oocita. Abbiamo contato, in media, 2000 nuclei in ciascun oocita! A quell'epoca era noto solo un altro vertebrato, la rana caudata (*Ascaphus truei*) dell'America Settentrionale, che avesse più di un nucleo per oocita: tuttavia ne possedeva appena otto. A quale scopo potrebbero servire tanti nuclei?

Una possibile spiegazione sta nel rapido sviluppo degli embrioni di *Flectonotus*. Essi vengono incubati per soli 20 giorni e i girini raggiungono la maturità dopo altri 11 giorni. Questo sviluppo accelerato probabilmente richiede una cospicua riserva di r-RNA. È presumibile che *Flectonotus* discenda da una specie

Il livello di RNA ribosomale (*in azzurro*) nell'oocita è correlato alla velocità dello sviluppo embrionale. In *Xenopus*, un anuro che depone le uova nell'acqua, l'r-RNA è abbondante per l'amplificazione dei geni ribosomali (*in rosso*) e l'embrione, molto vulnerabile, ha un rapido sviluppo: sette ore dopo la fecondazione si è già diviso in circa 4000 cellule. Nella raganella marsupiale *Gastrotheca*, come nei mammiferi (per esempio il topo), lo sviluppo può avvenire lentamente perché l'embrione è protetto. Il livello di r-RNA è basso e rispecchia una amplificazione genica scarsa o nulla: dopo sette ore l'embrione è ancora formato da una sola cellula o al massimo due. In certe raganelle che incubano le uova, come *Flectonotus*, l'oocita è multinucleato, il che aumenta il livello di r-RNA e accelera lo sviluppo.

in grado di incubare le uova, nella quale lo sviluppo embrionale risultava rallentato e l'amplificazione dei geni ribosomali era ridotta. Uno sviluppo più rapido potrebbe essere stato favorito dalle piccole dimensioni di *Flectonotus* e dal suo breve ciclo vitale. L'introduzione in ciascun oocita di migliaia di nuclei, ognuno dotato di una limitata amplificazione dei geni ribosomali, sarebbe la risposta evolutiva alla necessità di una maggior quantità di r-RNA. Tale strategia è risultata efficace, come Herbert C. Macgregor dell'Università di Leicester e io abbiamo scoperto misurando il livello totale di amplificazione genica.

Abbiamo poi scoperto che, per quanto siano insoliti in altri organismi, gli oociti multinucleati sono frequenti nelle raganelle che incubano le uova: li abbiamo trovati in 14 delle 36 specie da noi esaminate. In ogni caso l'adattamento si è presumibilmente evoluto in risposta a una pressione ambientale che favoriva uno sviluppo embrionale più rapido. Ma questo carattere solleva numerosi problemi. Per esempio, a mano a mano che l'oocita multinucleato si sviluppa nell'ovaio, il numero di nuclei diminuisce precipitosamente, finché ne rimane uno solo che si unisce al nucleo dello spermatozoo durante la fecondazione. Come viene controllata questa eliminazione di nuclei? Come può essere risparmiato un solo nucleo su migliaia?

Che si sviluppi da un oocita uninucleato o da uno multinucleato, l'uovo di una raganella marsupiale è caratteristico dal punto di vista strutturale e molecolare. È prevedibile che queste differenze influenzino l'andamento dello sviluppo embrionale, la complessa coreografia di crescita e movimento cellulare che inizia non appena l'uovo viene fecondato e trasportato nel marsupio. Richard P. Elinson dell'Università di Toronto e io abbiamo studiato le prime fasi dello sviluppo embrionale di *Gastrotheca*. Anche in questo caso abbiamo notato differenze rispetto all'andamento normale dello sviluppo negli anuri e rassomiglianze con quello di altri animali, specialmente uccelli e rettili.

Notevoli anomalie appaiono alla gastrulazione, quando l'embrione primitivo si ripiega verso l'interno e le cellule di un lato della sfera migrano per coprire la cavità risultante, che costituisce l'intestino primitivo. Nella maggior parte degli embrioni di anuro l'intestino primitivo si estende per quasi tutta la gastrula. Perciò il piano corporeo originale e lo sviluppo embrionale successivo coinvolgono l'intero embrione. Invece negli uccelli e nei rettili la grande quantità di tuorlo presente nelle prime fasi intralcia la gastrulazione. Solo un piccolo disco di cellule concentrate su un lato dell'uovo prende parte ai movimenti che formano l'intestino primitivo e solo queste cellule procedono alla formazione dell'organismo. Il resto dell'embrione è costituito da una massa grossa e indifferenziata di tuorlo, che viene consumata a poco a poco dall'embrione in via di sviluppo.

Elinson e io abbiamo scoperto che nelle raganelle che incubano le uova, come negli altri anuri, l'intero embrione prende parte alla gastrulazione: a partire da un'estremità, piccole cellule con poco tuorlo si ridispongono e formano un nuovo strato di cellule che va a coprire l'intera massa. Tuttavia l'invaginazione che normalmente avviene durante la gastrulazione è ridotta e l'intestino primitivo che si forma rimane limitato a una piccola cavità situata a un'estremità dell'embrione. Solo il disco di cellule che circonda questo piccolo intestino primitivo partecipa direttamente all'ulteriore sviluppo embrionale. L'embrione in via di sviluppo è adagiato su un lato di una grande massa di tuorlo: il modello di sviluppo embrionale ricorda perciò quello degli uccelli e dei rettili.

Centinaia di milioni di anni di evoluzione separano le raganelle marsupiali dagli animali prettamente terricoli - mammiferi, uccelli e rettili - con cui le ho messe a confronto. Eppure queste raganelle hanno sviluppato molte risposte analoghe alla sfida adattativa costituita dalla riproduzione sulla terraferma. A questo punto, si possono porre domande stimolanti sul potenziale evolutivo delle raganelle marsupiali: potrebbero esse con l'andar del tempo dare origine a una classe completamente nuova di vertebrati terrestri? Non sarebbe la prima conquista della terraferma da parte degli anfibi: i lontani progenitori dei mammiferi, degli uccelli e dei rettili possono essere identificati proprio in forme ancestrali di anfibi.

Tuttavia c'è la minaccia che per molte specie di raganelle marsupiali l'avventura evolutiva stia purtroppo volgendo al termine. Se non verranno presi efficaci e tempestivi provvedimenti si assisterà alla scomparsa della foresta tropicale, il solo habitat di questi animali; come conseguenza, anche molte singolari creature che ne abitano il baldacchino saranno destinate a una rapida estinzione. Le specie meno note di raganelle marsupiali potrebbero essere completamente spazzate via prima che si riescano a svelare tutti i loro segreti.

BIBLIOGRAFIA

DEL PINO EUGENIA M. e HUMPHRIES A. A., JR., *Multiple Nuclei During Early Oogenesis in* Flectonotus pygmaeus *and Other Marsupial Frogs* in «The Biological Bulletin», 154, n. 2, aprile 1978.

DEL PINO EUGENIA M. ed ELINSON RICHARD P., *A Novel Development Pattern for Frogs: Gastrulation Produces an Embryonic Disk* in «Nature», 306, n. 5943, 8 dicembre 1983.

DEL PINO E. M., STEINBEISSER H., HOFMANN A., DREYER C., CAMPOS M. e TRENDELENBURG M. F., *Oogenesis in the Egg-Brooding Frog* Gastrotheca riobambae *Produces Large Oocytes with Fewer Nucleoli and Low RNA Content in Comparison to* Xenopus laevis in «Differentiation», 32, pp. 24-33, 1986.

Le applicazioni mediche delle fibre ottiche

Le tecniche diagnostiche e chirurgiche che, mediante le fibre ottiche, consentono di operare in modo incruento su organi interni progrediscono costantemente e presto saranno impiegate nei casi di tumore e aterosclerosi

di Abraham Katzir

Se da una parte le fibre ottiche stanno trasformando la tecnica delle comunicazioni, dall'altra stanno rivoluzionando gli strumenti medici di indagine nonché i metodi di diagnosi e di terapia. Queste fibre ultrafini e flessibili hanno spalancato una finestra su organi e apparati del corpo umano. Introducendo fibre ottiche attraverso gli orifizi naturali o attraverso piccole incisioni e facendole scorrere lungo le vie dell'organismo, lo sguardo dello specialista può raggiungere le vie aeree inferiori, le anse intestinali, le cavità cardiache e molte altre zone in precedenza inaccessibili. Immergendo sensori a fibre ottiche nel flusso ematico, il medico può eseguire analisi chimiche rapide e sicure al letto del paziente, in ambulatorio o in sala operatoria, analisi che altrimenti comporterebbero un prelievo di sangue e l'invio del campione in laboratorio. Inviando un fascio di luce laser lungo una fibra ottica (detta anche fibra di potenza per distinguerla dalle fibre del fascio di visione), il chirurgo può addirittura eseguire un'operazione chirurgica all'interno dell'organismo evitando, quando possibile, il ricorso a tecniche invasive che comportino l'incisione di tessuti sani per raggiungere la parte malata.

Per esempio, già oggi è possibile eseguire delicati interventi inviando un fascio di luce laser tramite fibre ottiche. In gastroenterologia si possono cauterizzare i vasi per interrompere un'emorragia intestinale, in chirurgia vascolare si cominciano a eseguire interventi di vaporizzazione sulle placche aterosclerotiche e sui trombi delle arterie periferiche e in neurochirurgia si potranno presto connettere le fibre nervose lesionate. È probabile che, tra non molto, i dispositivi a fibre ottiche consentano di combinare diagnosi e terapia avendo incorporati, per esempio, un mezzo per riconoscere le cellule cancerose e uno per distruggerle senza danneggiare il tessuto sano circostante.

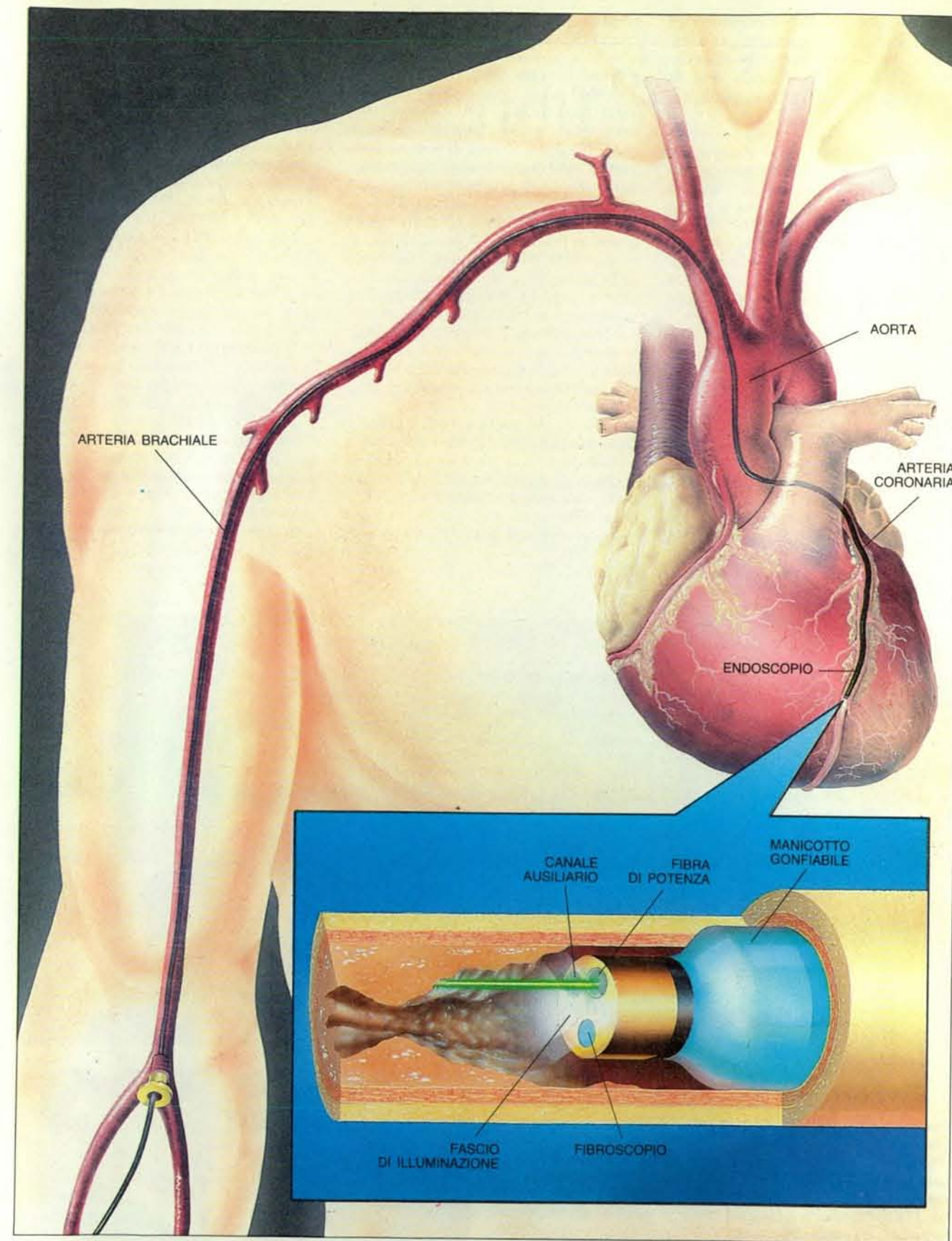
Molte tecniche diagnostiche e terapeutiche basate sulle fibre ottiche non richiedono anestesia e possono essere eseguite in tutta sicurezza in ambulatorio; il loro sviluppo dovrebbe quindi via via ridurre i rischi e i costi delle cure mediche. Un altro vantaggio delle tecniche basate sulle fibre ottiche è che possono essere applicate anche nei casi in cui la chirurgia tradizionale è rischiosa o impossibile, come accade talora in pazienti giovanissimi o molto anziani.

Le prime applicazioni mediche delle fibre ottiche furono i fibroscopi, strumenti di indagine visiva. Il primo fibroscopio fu costruito nel 1957 da Basil I. Hirschowitz e Lawrence Curtis, della School of Medicine dell'Università del Michigan, allo scopo di esaminare la cavità interna dello stomaco e dell'esofago. Da allora questo dispositivo è stato perfezionato a tal punto che oggi lo si può impiegare per esaminare in pratica qualunque organo o apparato. In effetti, le fibre ottiche impiegate in medicina sono per lo più incorporate nei fibroscopi. Un fibroscopio moderno contiene due fasci di fibre ottiche. Uno, il fascio di illuminazione, porta la luce ai tessuti e l'altro, il fascio di visione, trasmette l'immagine all'osservatore.

Il fascio d'illuminazione è accoppiato con un'intensa sorgente luminosa, per esempio una lampada ad arco allo xeno. La luce penetra nel nucleo delle fibre ottiche che sono fatte di vetro silicatico purissimo. Questo vetro è diecimila volte più trasparente del comune vetro da finestra e può quindi trasmettere la luce per molti chilometri. Poiché la luce tende a disperdersi uscendo dal nucleo, questo è avvolto da un mantello che riflette la maggior parte della luce rinviandola all'interno della fibra. Questo è il meccanismo con il quale la luce si propaga in tutte le fibre ottiche (si veda l'articolo *Le comunicazioni su fibre ottiche* di J. S. Cook in «Le Scienze» n. 66, febbraio 1974).

La luce riflessa dal tessuto bersaglio è raccolta da un obiettivo che la concentra sull'estremità ricevente del fascio di visione. Ciascuna fibra ottica di questo fascio può ricevere solo luce riflessa in linea con l'asse e quindi ogni fibra trasmette solo una piccola frazione dell'immagine complessiva. Le fibre sono saldate tra loro solo alle estremità il che, da un lato, conferisce flessibilità al fascio, mentre, dall'altro, evita che l'immagine giunga disordinata. L'immagine ricostruita può essere osservata attraverso un oculare, registrata su pellicola fotografica oppure inviata a uno schermo televisivo. Poiché in un unico fascio del diametro di meno di un millimetro si possono combinare migliaia di fibre, un fibroscopio può trasmettere immagini con elevata risoluzione spaziale e cromaticità quasi perfetta.

I fasci d'illuminazione e di visione possono essere facilmente alloggiati in un catetere di pochi millimetri di diametro. Le dimensioni ridotte dello strumento fanno sì che esso possa venire introdotto nel corpo attraverso un orifizio naturale e che possa essere diretto su tessuti distanti da cinque a cento millimetri dalla sua estremità.

Spesso i fibroscopi sono inseriti in strumenti più grandi, gli endoscopi, dotati di canali ausiliari che consentono allo specialista di ampliare le prestazioni dello strumento. Un canale può servire, per esempio, ad aspirare liquidi o a iniettare acqua o aria per allontanare detriti e migliorare la visibilità. Un altro canale può contenere sottili fili che vengono tirati per orientare l'estremità dell'endoscopio. Un terzo canale può consentire l'introduzione di minuscoli bisturi per ta-

Utilizzando le fibre ottiche si possono inviare fasci di luce laser nelle coronarie per vaporizzare le placche aterosclerotiche. Uno dei dispositivi realizzabili in un prossimo futuro comprende un fibroscopio, un manicotto gonfiabile e una fibra per il trasporto della luce laser. Questo dispositivo potrebbe essere introdotto in un'arteria del braccio ed essere spinto fino al cuore. Il fibroscopio consentirebbe al medico di vedere le placche o altre ostruzioni. Gonfiando il manicotto il flusso ematico verrebbe temporaneamente arrestato; il fascio laser inviato attraverso la fibra farebbe vaporizzare la placca. Sgonfiando il manicotto, si ripristinerebbe la circolazione.

gliare i tessuti o di pinze per rimuoverli o può essere dotato di siringhe per iniettare farmaci. La lunghezza degli endoscopi attualmente disponibili varia da 0,3 a 1,2 metri e il loro diametro da 2,5 a 15 millimetri.

Con questi strumenti i medici possono osservare da vicino l'interno degli apparati digerente, riproduttivo, circolatorio e respiratorio. Possono rimuovere frammenti di tessuto da sottoporre a esame bioptico ed effettuare addirittura operazioni chirurgiche. Attraverso l'endoscopio si possono scoprire un polipo nel colon, un corpo estraneo nei polmoni o un tumore nell'esofago e poi asportarli con strumenti chirurgici in miniatura.

Negli ultimi cinque anni la fabbricazione di fibre ultrafini ha permesso di ridurre il diametro dei fibroscopi e di aumentare il numero di fibre nel fascio di visione, il che ha consentito un aumento di risoluzione. Gli strumenti più recenti incorporano 10 000 fibre in un fascio di diametro inferiore a un millimetro e la loro risoluzione raggiunge i 70 micrometri. Introdotti in un'arteria del braccio, questi fibroscopi possono fornire immagini delle valvole cardiache e di eventuali ostruzioni nelle coronarie, le arterie che alimentano il cuore.

Oltre a fornire immagini, i sistemi a fibre ottiche possono offrire un'analisi chimica e fisica diretta e immediata del sangue e di altri elementi importanti per la fisiologia umana. Il sensore è costituito fondamentalmente da una fibra ottica introdotta nel corpo attraverso un catetere. L'estremità esterna della fibra è accoppiata con una sorgente luminosa e con un processore in grado di analizzare la luce riflessa. L'altra estremità della fibra invia la luce direttamente dentro la cavità da esaminare oppure a un sensore in miniatura detto optodo. La luce riflessa ripercorre all'indietro la fibra fino al processore che, dalla lunghezza d'onda e intensità della luce, ricava informazioni sulle condizioni fisiologiche del paziente.

In molti casi questi sistemi a fibra ottica sono più sensibili, affidabili ed economici dei metodi tradizionali basati sul prelievo di liquidi organici e sulla loro successiva analisi in laboratorio. I sistemi a fibra ottica abbreviano i tempi e riducono le possibilità d'errore. Inoltre i sensori a fibra ottica sono biocompatibili (per esempio non provocano reazioni del sistema immunitario) e, in linea di principio, la stessa fibra ottica può non essere recuperata. I sistemi a fibra ottica, inoltre, appaiono più resistenti, flessibili e potenzialmente meno rischiosi dei dispositivi microelettronici che sono stati allestiti per gli stessi compiti. Alcuni dei dispositivi a fibra ottica sopra descritti sono già sul mercato e molti altri sono sottoposti a sperimentazione clinica.

Le tecniche di misura del flusso ematico mediante fibre ottiche si basano sulla luce riflessa dai globuli rossi. Una fibra viene introdotta in un'arteria attraverso un catetere in modo che una luce laser di bassa intensità inviata attraverso la fibra colpisca i globuli rossi. Quando la luce viene diffusa dalle cellule in movimento, la sua lunghezza d'onda varia per effetto Doppler. Quanto più veloce è il moto della cellula verso l'estremità della fibra, tanto più breve è la lunghezza d'onda della luce diffusa. Una parte di questa luce torna indietro lungo la fibra. Il processore situato all'altra estremità della fibra può così determinare la differenza di lunghezza d'onda tra la luce laser e la luce diffusa e calcolare la velocità del sangue nella zona dove è immersa la fibra.

Il fibroscopio (*in alto*) è uno strumento che può trasmettere immagini dello stomaco e di molti altri organi. Una lente focalizza la luce emessa da una lampada su un fascio di fibre ottiche. In questo caso la luce trasmessa dalle fibre va a illuminare un polipo dello stomaco. Un obiettivo concentra la luce riflessa dal polipo su un fascio ordinato di fibre, ciascuna delle quali trasporta una frazione dell'immagine complessiva. Emettendo la luce all'estremità opposta, le fibre del fascio danno un'immagine ricostruita che può essere osservata mediante un oculare. Le fotografie del polipo inserite nell'illustrazione sono state ottenute mediante fibroscopi presso la Mayo Clinic. I fibroscopi sono spesso incorporati in dispositivi più grandi, gli endoscopi (*a destra*), dotati di canali attraverso i quali possono essere fatti passare altri strumenti. Qui a destra è rappresentata una porzione di tessuto bronchiale mentre viene rimossa mediante uno strumento inserito attraverso l'endoscopio.

I primi esperimenti che dimostrarono la validità di questa tecnica furono descritti quasi trent'anni fa, ma la sperimentazione clinica ha avuto inizio solo da pochi anni. Disponendo di misurazioni istantanee della velocità del sangue fornite da uno strumento altamente affidabile, i medici avranno la possibilità di stabilire se gli organi vitali sono irrorati a sufficienza.

Le fibre ottiche consentono anche di determinare direttamente il contenuto di ossigeno nel sangue. L'emoglobina (la molecola che trasporta l'ossigeno nel sangue) riflette di più la luce rossa quando contiene ossigeno che quando non lo contiene. Invece la luce infrarossa viene riflessa dell'emoglobina sempre nella stessa maniera, indipendentemente dal suo contenuto di ossigeno. Se mediante fibre ottiche vengono immesse nel sangue luce rossa e infrarossa l'intensità della luce rossa riflessa rivela la quantità di emoglobina che trasporta ossigeno, mentre l'intensità della luce infrarossa riflessa misura la quantità di emoglobina complessiva. Questa tecnica, oggi di routine, consente di valutare la capacità di trasporto dell'ossigeno attraverso il sangue del paziente e di stabilire con quale efficienza cuore e polmoni riforniscono il sangue di ossigeno.

Sensori in miniatura collegati alle estremità di fibre ottiche hanno reso possibili molte altre misurazioni di notevole importanza per lo studio della fisiologia umana. Sono stati ideati sensori per misurare la pressione all'interno delle arterie, della vescica, dell'uretra e del retto. Un sensore è costituito da un tubicino collocato all'estremità di una fibra ottica; l'estremità più lontana del tubicino è sigillata da una sottile membrana riflettente. Se la pressione interna al tubicino è uguale all'esterna, la membrana ha una forma piana e la luce trasportata dalla fibra viene semplicemente riflessa all'indietro. Se la pressione esterna è maggiore di quella interna, la membrana s'incurva verso l'interno e forma uno specchio convesso che riflette meno luce. Se viceversa prevale la pressione interna, la membrana si incurva verso l'esterno e la sua superficie concava convoglia più luce nella fibra. In parecchi esperimenti clinici è stata trovata una buona correlazione tra i risultati ottenuti con un dispositivo ottico di questo genere e quelli ottenuti con strumenti tradizionali.

Sono anche stati costruiti sensori che misurano la composizione chimica del sangue. Quelli allestiti per misurare il *p*H (cioè l'acidità o l'alcalinità del sangue) sono basati su coloranti organici che, in presenza di radiazione ultravioletta, diventano luminescenti. Se il *p*H si modifica, alcuni di questi coloranti diventano più luminescenti mentre altri emettono una luminescenza di colore diverso. Questi coloranti, inglobati in un polimero poroso, vengono fissati all'estremità di una fibra. Il polimero consente l'ingresso agli ioni idrogeno, ma impedisce che ioni più grandi interagiscano con il colorante e lo degradino. In pratica questo sensore può misurare il *p*H del sangue con una precisione di 0,01 unità.

Si possono anche produrre sensori di concezione analoga che però incorporano biomolecole come enzimi o anticorpi. Una modifica della fluorescenza di queste molecole può rivelare la presenza o l'assenza di particolari sostanze chimiche in una miscela piuttosto complessa com'è il sangue o in un tessuto biologico. Anche in questo caso la fibra ottica trasmette le informazioni dal sensore a un processore che interpreta le informazioni e fornisce misurazioni estremamente precise di un'ampia varietà di sostanze chimiche presenti all'interno del corpo. Questi sensori sono già stati impiegati per seguire le variazioni nella concentrazione del glucosio e della penicillina e saranno presto impiegati per misurare i livelli di farmaci, metaboliti, ormoni, tossine e microrganismi.

Il ventricolo destro di un cuore umano è stato fotografato attraverso un fibroscopio ultrasottile inserito in un'arteria del braccio. Il fibroscopio, avente diametro inferiore a un millimetro, è una realizzazione della Olympus Corporation di Tokyo.

Negli ultimi anni l'applicazione più significativa delle fibre ottiche in medicina ha riguardato la terapia e la chirurgia laser. Le modalità d'interazione fra radiazione laser e tessuti umani dipendono dalla lunghezza d'onda e dall'intensità della radiazione. Benché la luce emessa da un laser particolare abbia un'unica lunghezza d'onda, o colore, oggi esistono svariati laser che generano luce in tutto lo spettro visibile, infrarosso e ultravioletto. Il livello di assorbimento della luce da parte dei tessuti è variabile e dipende dalla lunghezza d'onda e dai cromofori contenuti nel tessuto come emoglobina, melanina e cheratina. Di conseguenza un laser di una data lunghezza d'onda ha come bersaglio tessuti specifici e produce reazioni fotochimiche specifiche.

In generale i laser di bassa potenza provocano un debole riscaldamento locale che fa coagulare sangue e proteine. In tal modo la luce del laser può saldare tra loro i tessuti molli e può quindi far rimarginare le ferite o chiudere i vasi sanguigni. I laser di elevata potenza vengono usati per l'ablazione dei tessuti e, nella maggior parte dei casi, agiscono facendo evaporare l'acqua in essi contenuta. Un fascio del genere cauterizza anche le incisioni limitando al massimo le emorragie durante le operazioni. Le applicazioni chirurgiche del laser richiedono una potenza da 10 a 100 watt nei laser continui o picchi di potenza da 10 000 a un milione di watt nei laser a impulsi. Dato che questa potenza viene concentrata su superfici di area inferiore al millimetro quadrato, le densità di potenza sono più o meno pari a quelle di un laser da «guerre stellari», in grado di perforare l'involucro di un missile guidato in fase di lancio.

Negli ultimi vent'anni i ricercatori si sono cimentati con le difficoltà connesse all'immissione nell'organismo di potenze così elevate attraverso le fibre ottiche. Una caratteristica di queste fibre è quella di attenuare in qualche misura la luce per assorbimento e diffusione. L'attenuazione dipende in parte dalle proprietà intrinseche del materiale della fibra; a sua volta, questa attenuazione intrinseca varia con la lunghezza d'onda della luce e con la potenza totale trasmessa. L'energia del raggio laser può essere attenuata anche a causa della diffusione sulla superficie della fibra o sui difetti interni di questa. Tutte queste cause limitano la potenza che può essere applicata sul punto dell'intervento. Se la potenza del fascio laser in ingresso viene aumentata oltre un livello critico, l'estremità della fibra può surriscaldarsi, fondere o vaporizzare. Negli ultimi anni molte di queste difficoltà sono state superate grazie alla scoperta di nuovi materiali e alla messa a punto di metodi per fabbricare con essi fibre di elevata purezza.

A differenza delle fibre di vetro silicatico usate nei sistemi diagnostici e di indagine visiva, le fibre per la chirurgia laser sono fatte di materiali alquanto ricercati. Si sono costruite fibre di quarzo che trasmettono la luce verde dei laser ad argo, la luce ultravioletta dei laser a eccimeri e la radiazione quasi infrarossa dei laser YAG (in cui il materiale attivo è un cristallo di granato contenente ittrio e alluminio e «drogato» con ioni neodimio). Tali fibre sono riuscite a trasmettere fasci laser aventi una potenza continua di 100 watt, sufficiente per la maggior parte degli interventi chirurgici. I ricercatori, tuttavia, sono ancora alla ricerca di una fibra capace di trasmettere efficientemente la radiazione dei laser infrarossi, in particolare dei laser ad anidride carbonica (si veda l'articolo *Fibre*

Un sensore a fibra ottica può misurare la pressione in vasi e cavità, come arterie, vescica e utero. Il sensore è costituito da un tubicino posto all'estremità di una fibra e chiuso da una membrana riflettente. Se la pressione esterna è superiore a quella interna (*a sinistra*), la membrana s'incurva verso l'interno, creando uno specchio convesso che riflette poca luce attraverso la fibra. Se la pressione interna prevale su quella esterna (*a destra*), la membrana s'incurva verso l'esterno e la sua superficie concava riflette più luce nella fibra.

ottiche per l'infrarosso di Martin G. Drexhage e Cornelius T. Moynihan in «Le Scienze» n. 245, gennaio 1989). Per il momento le migliori fibre ottiche nell'infrarosso sono fibre policristalline costituite da alogenuri metallici.

Il primo dispositivo che utilizzava il laser e le fibre ottiche in campo medico fu progettato nel 1973 per arrestare le emorragie delle ulcere peptiche. Questo strumento consisteva in un fibroscopio con cui osservare il procedimento e in una fibra ottica separata che trasportava l'energia laser per cauterizzare l'ulcera. Da allora questi dispositivi sono stati impiegati con successo per intervenire su ulcere emorragiche dello stomaco, dell'intestino e del colon. Una fibra ottica per il trasporto di luce laser è stata impiegata anche per frantumare i calcoli renali. Le prospettive terapeutiche più interessanti per i sistemi a fibre laser riguardano tuttavia la cura delle malattie cardiovascolari e dei tumori localizzati.

Molte delle malattie cardiovascolari più letali insorgono quando le arterie vengono bloccate da depositi di grassi calcificati e fibrosi, le placche aterosclerotiche, e da trombi sanguigni. Se queste ostruzioni arrivano a interrompere la circolazione, provocano ictus, infarti e cancrena degli arti. Per curare un paziente le cui arterie coronarie siano ostruite, lo specialista può oggi ricorrere a un metodo chiamato angioplastica coronarica transluminale percutanea. Questa tecnica fa uso di un catetere con un palloncino attaccato all'estremità. Se un'arteria è ostruita solo in parte, l'estremità del catetere può essere inserita nella regione dove il lume è ridotto. Il palloncino viene gonfiato in modo da comprimere e ridurre l'occlusione. Questa tecnica non va tuttavia applicata quando l'occlusione è totale e i suoi risultati positivi sono a volte solo temporanei.

Il cardiochirurgo può allora raccomandare un intervento più radicale, il *bypass* coronarico; esso consiste nel prelevare una vena dalla gamba e nell'inserirla nel cuore dopo aver aperto il torace. La vena serve a trasportare il sangue oltre l'occlusione arteriosa e, di solito, riesce a ripristinare l'irrorazione del cuore; tuttavia, si tratta di un procedimento traumatico che comporta una lunga convalescenza e spese rilevanti.

La realizzazione di fibre ottiche capaci di trasportare un fascio laser di elevata potenza ha reso possibili parecchie nuove tecniche per rimuovere le ostruzioni arteriose, tecniche che vanno sotto il nome di angioplastica laser. Alcune di queste tecniche utilizzano una fibra la cui estremità è coperta da un cappuccio metallico. Se la fibra viene inserita nell'arteria ostruita e attraverso di essa viene trasmessa luce laser, il cappuccio metallico si scalda, disintegrando l'ostruzione. Questa tecnica richiede un controllo accuratissimo, per evitare che il cappuccio riscaldato si attacchi alla parete arteriosa o addirittura la perfori. Oggi, per l'angioplastica laser, disponiamo di parecchi, validi strumenti.

Un'alternativa potenzialmente più efficace, ma tecnicamente più complessa è rappresentata da un sistema in cui la luce laser asporta direttamente la placca. I primi sistemi sperimentali di questo tipo utilizzavano un laser ad argo, che emette luce verde. Questo laser era stato scelto per la sua affidabilità e perché la sua luce poteva essere trasmessa in modo efficiente dalle tradizionali fibre di quarzo. Gli esperimenti misero tuttavia in evidenza che la luce verde assorbita provocava estesi danni termici ai tessuti circostanti e non asportava bene la placca. Questi problemi poterono essere superati usando un fascio laser a impulsi che emetteva nell'ultravioletto o nell'infrarosso. Di recente Warren S. Grundfest, James S. Forrester e Frank Litvack del Cedars-Sinai Medical Center di Los Angeles hanno collaudato un sistema basato su un laser a eccimeri che emette nell'ultravioletto e su una fibra di quarzo. Il loro sistema è riuscito ad eliminare le ostruzioni delle arterie coronarie in parecchi pazienti.

I problemi più seri ancora da risolvere consistono nel mettere a punto un sistema per manovrare il fascio laser dentro l'arteria e nel trovare un modo per distinguere i segmenti di arteria sani da quelli malati prima di usare il laser per vaporizzare i tessuti. In futuro vi saranno sistemi «intelligenti», basati sulla fluorescenza endoscopica, capaci di eseguire queste operazioni senza il pericolo di perforare i vasi sanguigni.

I miei colleghi dell'Università di Tel Aviv e io crediamo che, nel prossimo futuro, nuove fibre ottiche capaci di trasmettere la radiazione infrarossa di un laser ad anidride carbonica consentiranno di allestire un sistema più sicuro e affidabile. Abbiamo collaudato questo sistema trapiantando in cavie arterie umane ostruite da placche aterosclerotiche. Abbiamo poi introdotto nelle arterie fibre ottiche che trasportavano la luce di un laser ad anidride carbonica per eliminare le ostruzioni.

Allo stato attuale si può prevedere che entro pochi anni tali strumenti chirurgici saranno perfezionati e incorporati in un dispositivo comprendente fibroscopio e sensori. Questo endoscopio laser avrà probabilmente un diametro inferiore a due millimetri e circa metà dello spazio

Le radiografie delle arterie femorali eseguite prima e dopo un intervento di angioplastica laser e applicazione di catetere a palloncino confermano l'efficacia dei trattamenti. Nella radiografia a sinistra l'arteria è del tutto ostruita. Dopo l'ablazione dell'ostruzione con il laser, resta solo il 30 per cento dell'occlusione primitiva (*al centro*). Per eliminare anche questa si può inserire nell'arteria un catetere a palloncino. Quando il catetere viene rimosso, la radiografia dimostra che il vaso è divenuto del tutto pervio (*a destra*).

Un tumore che ostruisce la trachea (*a sinistra*) viene trattato con la terapia fotodinamica. Si inietta un colorante che viene assorbito più facilmente dai tessuti tumorali che da quelli sani. Una fibra ottica inserita nel tumore (*al centro*) trasmette luce laser che modifica chimicamente il colorante. Due giorni dopo (*a destra*), il colorante modificato ha distrutto il tumore rendendolo asportabile.

sarà occupato dal fibroscopio. Il dispositivo conterrà anche una fibra ottica per la trasmissione di luce laser ultravioletta, infrarossa o visibile. Sempre nell'endoscopio saranno alloggiati sensori a fibre ottiche aventi lo scopo di misurare la pressione, la temperatura e il flusso del sangue, mentre attraverso un altro canale sarà possibile pompare liquidi e gas.

L'endoscopio potrebbe essere introdotto nell'arteria coronaria come un catetere ordinario. Lo specialista potrebbe così agevolmente osservare l'ostruzione e misurare la pressione e il flusso del sangue. Il flusso potrebbe essere arrestato con un manicotto gonfiabile, mentre soluzione fisiologica trasparente verrebbe iniettata attraverso il canale d'irrigazione allo scopo di asportare il sangue residuo. A questo punto attraverso la fibra di potenza si potrebbe inviare un fascio di luce laser per vaporizzare l'ostacolo. I gas sviluppati dalla vaporizzazione della placca verrebbero eliminati attraverso il canale d'aspirazione oppure smaltiti dall'organismo. I sensori a fibra ottica, collegati a uno speciale processore, seguirebbero le fasi dell'intervento segnalando il surriscaldamento dell'estremità della fibra o l'ablazione di un vaso sano. Al termine dell'angioplastica il manicotto gonfiabile verrebbe sgonfiato e il flusso sanguigno misurato per confermare che l'arteria coronaria è stata riaperta.

Un'altra applicazione in via di sviluppo è l'impiego delle fibre ottiche per la diagnosi e la cura di piccoli tumori maligni. Il metodo diagnostico dell'endoscopia a fluorescenza ha avuto un notevole successo nella diagnosi di tumori polmonari troppo piccoli per essere individuati dalla tomografia assiale computerizzata (TAC) o dalla radiografia toracica. Si è osservato che se nel paziente viene iniettata una sostanza nota come derivato dell'ematoporfirina, un colorante che emette fluorescenza rossa quando è colpito da radiazione ultravioletta, dopo alcuni giorni i tumori mostrano di aver assorbito molto più colorante dei tessuti sani. Se la regione sospetta viene illuminata con una sorgente ultravioletta idonea, per esempio una fibra di quarzo accoppiata con un laser al cripto, il tessuto maligno diventa riconoscibile in quanto emette luce rossa. Per rilevarla, si collega al fascio di visione un filtro che trasmette la luce rossa e blocca la luce ultravioletta assorbita.

Se il tessuto viene colpito da luce rossa, anziché ultravioletta, abbastanza intensa, i risultati sono profondamente diversi. Il derivato dell'ematoporfirina assorbe fortemente la luce rossa e l'energia assorbita provoca una serie di reazioni fotochimiche che uccidono i tessuti maligni in precedenza saturati con derivato dell'ematoporfirina. Un laser a vapori d'oro può fornire la luce rossa a elevata intensità che viene trasmessa attraverso una fibra di quarzo direttamente all'interno del tumore. Questa luce dunque distrugge selettivamente le cellule cancerose. Tale terapia fotodinamica è attualmente in fase di sperimentazione clinica (per approfondire l'argomento si veda l'articolo *Il laser in oncologia* di C. A. Sacchi e P. Spinelli in *Progressi nella ricerca sul cancro*, Le Scienze, 1989).

Nel prossimo futuro questi sistemi diagnostici e chirurgici potrebbero essere abbinati in un unico endoscopio laser per la cura dei tumori. Il dispositivo comprenderebbe un fibroscopio, una fibra di quarzo per la trasmissione della luce d'eccitazione ultravioletta e un'altra fibra di quarzo per la trasmissione della luce rossa per la terapia fotodinamica. L'endoscopio verrebbe introdotto attraverso un orifizio naturale o attraverso la cute e raggiungerebbe il sospetto tumore. La luce laser ultravioletta sarebbe trasmessa attraverso la prima fibra e l'immagine fornita dal fibroscopio verrebbe osservata attraverso un filtro che lascia passare selettivamente il rosso. Le neoplasie sarebbero identificate grazie alla loro emissione rossa e, attraverso la seconda fibra di quarzo, verrebbe inviata luce rossa di intensità elevata per distruggerle. In caso di tumore di grandi dimensioni, la fibra ottica potrebbe essere inserita direttamente dentro la neoformazione. Dopo alcune settimane (o dopo alcune applicazioni) il procedimento potrebbe essere ripetuto per sincerarsi che la crescita del tumore si sia arrestata.

Fibroscopi flessibili, sensori ottici e sistemi laser aprono un'era di tecniche mediche pochissimo invasive. Gli endoscopi laser comprenderanno un fibroscopio completo di fascio di visione e di guide per l'illuminazione, fibre ottiche accoppiate con sensori diagnostici, una fibra di potenza per erogare la radiazione laser e un canale per iniettare liquidi e aspirare liquidi o gas. Lo specialista disporrà così di uno strumento dotato contemporaneamente della chiarezza di un televisore ad alta definizione, della precisione di un laboratorio medico, dell'abilità di un'équipe di chirurghi e di dimensioni che gli consentono di passare attraverso le più piccole cavità del corpo umano.

BIBLIOGRAFIA

ALLEN W. B., *Fiber Optics: Theory and Practice*, Plenum Press, 1973.

WOLF H. F. (a cura), *Handbook of Fiber Optics: Theory and Applications*, Granada Publishing, Inc., 1979.

DIXON JOHN A., *Lasers in Surgery* in «Current Problems in Surgery», 21, n. 9, settembre 1984.

Optical Fibers in Medicine IV in «Proceedings of the Society of Photo-Optical Instrumentation Engineers», a cura di A. Katzir, 1067, giugno 1989.

Henry Norris Russell (1877-1957), «decano degli astronomi statunitensi», nacque a Oyster Bay (New York) e trascorse la maggior parte della vita a Princeton. Sono legate al suo nome le scoperte del diagramma di Hertzsprung-Russell, che correla la luminosità intrinseca delle stelle con il loro colore, e dell'accoppiamento di Russell-Saunders, che descrive le interazioni di due elettroni negli spettri atomici. Dopo la pubblicazione, nel 1920, di un articolo del fisico indiano Meghnad N. Saha in cui si presentava una teoria della ionizzazione fondata sulla meccanica quantistica, Russell si dedicò al tentativo di spiegare razionalmente gli spettri stellari.

Henry Norris Russell

Fu uno dei maggiori astronomi della sua generazione: convinto della necessità di fornire all'astronomia una solida base teorica, perseguì tale obiettivo contribuendo così alla fondazione dell'astrofisica moderna

di David H. DeVorkin

Nel luglio 1923 George Ellery Hale ricevette da un astronomo del Mount Wilson Observatory, di cui era direttore, la seguente relazione su quello che era diventato un piacevole rituale estivo: «Al suo arrivo Henry Norris Russell si mise subito al lavoro di buona lena... I colloqui si trasformarono ben presto in un assolo che continuò senza pause per tutto il suo soggiorno. Ci riunivamo tre o quattro volte la settimana, ed egli ci parlava delle serie spettrali applicate alla struttura atomica... Negli ultimi tempi, prima di ripartire, dedicò la maggior parte del suo tempo a sviluppare la serie del titanio... La complessità del problema si è rivelata maggiore di quanto egli si attendesse e il lavoro non è ancora terminato, ma la selezione fondamentale è stata fatta ed è stato chiarito il carattere dei multipletti...».

Questo fu Henry Norris Russell nella pienezza del suo vigore intellettuale: una sorgente continua di idee su come incorporare la fisica moderna nell'astronomia spettroscopica. Fra i suoi meriti si annoverano il diagramma di Hertzsprung-Russell, il metodo di Russell per le binarie a eclisse, l'accoppiamento di Russell-Saunders e l'opera fondamentale da lui scritta insieme a Raymond S. Dugan e J. Q. Stewart, *Astronomy: Analysis of Stellar Spectra*, che contribuì a formare due generazioni di astronomi. E fu ancora lui che per più di 40 anni collaborò con una rubrica mensile di astronomia alle pagine di «Scientific American».

Russell, che aveva in sé una straordinaria carica di energia, fu inviato a Mount Wilson da Hale per utilizzare il patrimonio di informazioni ivi accumulate su spettri di laboratorio e spettri celesti e per insegnare agli astronomi dell'osservatorio a coniugare i loro studi osservativi con le potenzialità esplicative della fisica moderna. Hale sperava che Russell e altri fisici saldassero la frattura fra il banco del laboratorio e la cupola dell'osservatorio. Il fatto che Russell si sia proposto come compito primario il problema di come unire questi due distinti livelli ci offre una possibilità di comprendere a fondo non solo la sua vita professionale, ma anche il carattere di una scienza in transizione.

Nel 1889 il dodicenne Henry, figlio primogenito di un pastore protestante di Oyster Bay (Long Island) fu mandato a vivere con la zia materna a Princeton per poter trarre vantaggio dalle buone scuole della cittadina del New Jersey. Dalla madre, Eliza Norris, aveva ereditato una predisposizione per la matematica e per la soluzione di problemi difficili e un acuto senso del dovere. Armato di una memoria infallibile, a sedici anni si diplomò alla Princeton Preparatory School e si iscrisse alla locale università, dove studiò matematica e astronomia. Studente universitario, subì l'influenza di Charles A. Young, un pioniere della spettroscopia solare, e preparò una tesi sulla classificazione visiva degli spettri stellari. Mentre continuava a lavorare per il dottorato ebbe fra i suoi *tutors* (i docenti che seguivano più da vicino le sue ricerche) oltre a Young, il matematico Henry B. Fine e l'astronomo E. O. Lovett. Questa mescolanza di matematica e di osservazione lo condusse a preparare una tesi di dottorato che allora era di particolare attualità: uno studio matematico delle perturbazioni gravitazionali esercitate da Marte sull'orbita di un pianetino allora appena scoperto, Eros (Witt, Berlino, 1898); l'analisi del moto di Eros da parte di Russell avrebbe condotto alcuni decenni dopo (Spencer Jones, 1931) alla determinazione di un valore più esatto della distanza Terra-Sole.

Fu in questo periodo formativo che Russell evidenziò le proprie caratteristiche peculiari, soprattutto una capacità straordinaria di scoprire nuove tecniche di calcolo. Egli seppe trovare metodi per il calcolo della massa delle stelle di sistemi binari visuali e per determinare la densità di stelle variabili del tipo di Algol, che a quell'epoca veniva considerata dalla maggior parte degli astronomi come un tipo speciale di binaria a eclisse. Russell si rese conto ben presto che anche altri erano impegnati nella corsa alla soluzione degli stessi problemi: Hendrikus J. Zwiers lo precedette di circa due anni nella determinazione della tecnica per le binarie visuali e Alexander Roberts suggerì la tecnica per le stelle del tipo di Algol quasi contemporaneamente a Russell.

I primi casi di competizione in cui Russell si trovò coinvolto ebbero l'effetto di spronarlo ancora di più. Dopo aver completato la sua tesi nel 1900, ebbe un esaurimento nervoso per l'eccesso di lavoro a cui si era sottoposto e l'anno seguente recuperò le energie viaggiando in Francia. Successivamente si recò come borsista all'Università di Cambridge, dove il suo ambizioso programma di ricerche teoriche e osservative sulle parallassi trigonometriche di alcune stelle venne nuovamente interrotto da un grave disturbo, forse un collasso nervoso. Lasciando incompiuta la sua ricerca, nel 1905 Russell tornò a Princeton, ottenne una docenza all'università e impiegò i cinque anni successivi nel tentativo di stabilire una relazione fra vari caratteri osservati e calcolati delle stelle: luminosità intrinseca, colore, massa, densità e spettro.

Per questo studio poté usufruire ampiamente degli spettri stellari fornitigli da Edward C. Pickering, il direttore dello Harvard College Observatory, la cui legione di assistenti femminili - fra le quali spicca il nome di Annie J. Cannon - aveva raccolto la più grande collezione mondiale di spettri e luminosità apparenti di stelle. Pickering suggerì a Russell di paragonare le luminosità assolute, che avrebbe potuto ricavare dalle parallassi stellari determinate a Cambridge, con gli spettri delle stesse stelle nella collezione di Harvard.

Il risultato finale di questa ricerca fu il famoso diagramma di Hertzsprung-Russell. (Il danese Ejnar Hertzsprung vi ar-

rivò indipendentemente fra il 1908 e il 1910.) Questo diagramma, pubblicato da Russell nel 1914 (*si veda l'illustrazione in questa pagina*), mostra come quasi tutte le stelle possano essere classificate in base alla luminosità e al colore e ha svolto un ruolo cruciale nella nascita della teoria dell'evoluzione stellare.

Nei medesimi anni lo stesso Russell stava sviluppando una teoria dell'evoluzione stellare (la sua teoria delle stelle giganti e delle stelle nane) fondata su un'ipotesi precedente di Norman Lockyer. Secondo questa teoria, una stella inizia la sua esistenza sotto forma di una nube di gas molto estesa (una gigante rossa), che va gradualmente contraendosi e riscaldandosi per effetto della propria attrazione gravitazionale sino a raggiungere un punto critico in corrispondenza del quale non si comporta più come un gas perfetto. Essa poi si raffredda contraendosi ulteriormente in una stella nana densa e passa il resto della propria vita a raffreddarsi e contrarsi sino a sparire. Si deve tener presente che a quei tempi le reazioni nucleari erano sconosciute e che la teoria si fondava semplicemente sull'attrazione gravitazionale e sulla teoria cinetica del calore. Gli astronomi disponevano inoltre di pochi elementi per dimostrare che le stelle si comportano come gas perfetti. La connessione fra la fisica moderna e le stelle era ancora debole.

Gli anni 1911 e 1912 videro un ulteriore contributo di Russell: il primo metodo rapido ed efficace per ricavare i parametri orbitali e le caratteristiche fisiche delle binarie a eclisse, che doveva diventare noto come metodo di Russell per le binarie a eclisse.

Grazie alle sue ricerche di astronomia stellare, Russell fu nominato professore ordinario di astronomia a Princeton, e qui continuò a insegnare per il resto della vita. Benché evitasse la politica e non avesse fondato alcun istituto, all'epoca della prima guerra mondiale Russell era diventato un'autorità nell'astronomia statunitense, in grado di imporre programmi di ricerca e di dirigere la vita professionale di ricercatori molto oltre i confini di Princeton. Uno dei suoi primi studenti, Harlow Shapley, divenne nel 1921 direttore dello Harvard College Observatory. Hale, Edwin B. Frost e Otto Struve, succedutisi alla direzione dell'«Astrophysical Journal», si rivolgevano a Russell per avere consigli su quali fossero gli articoli accettabili per essere pubblicati sulla rivista.

Benché gli allievi di Russell siano stati pochi, fra di essi si contarono personalità di grande rilievo. Donald H. Menzel diresse tentativi mirati alla comprensione di ciò che avviene negli strati esterni del Sole e della fisica delle nebulose gassose e nel 1954 divenne direttore dello Harvard College Observatory. Lyman Spitzer, Jr., fu un pioniere nell'applicazione della fisica dei plasmi alle stelle e quando, nel 1947, Russell andò in pensione, divenne direttore del Princeton University Observatory. Di fatto, passando in rassegna il panorama dell'astronomia statunitense degli anni trenta, Struve giunse alla conclusione che l'unico posto in cui si potesse studiare astrofisica teorica era Princeton, presso Henry Norris Russell.

Il diagramma di Hertzsprung-Russell (H-R) fu sviluppato indipendentemente fra il 1908 e il 1913 dal danese Ejnar Hertzsprung e da Russell. Esso riporta in ordinata la luminosità, o magnitudine, assoluta di una stella, e in ascissa la classificazione spettrale, che va dalle stelle blu (O) alle stelle rosse (M). La maggior parte delle stelle si dispone nella sequenza principale, che corre diagonalmente da sinistra in alto a destra in basso. Le giganti rosse cadono in alto a destra e le nane bianche si collocano in basso a sinistra. L'astrofisica del XX secolo si è dedicata in gran parte al tentativo di spiegare teoricamente il diagramma H-R. Il diagramma presentato è la prima versione di Russell, pubblicata in «Popular Astronomy» nel 1914. I cerchietti indicano diversi livelli di confidenza per le parallassi stellari.

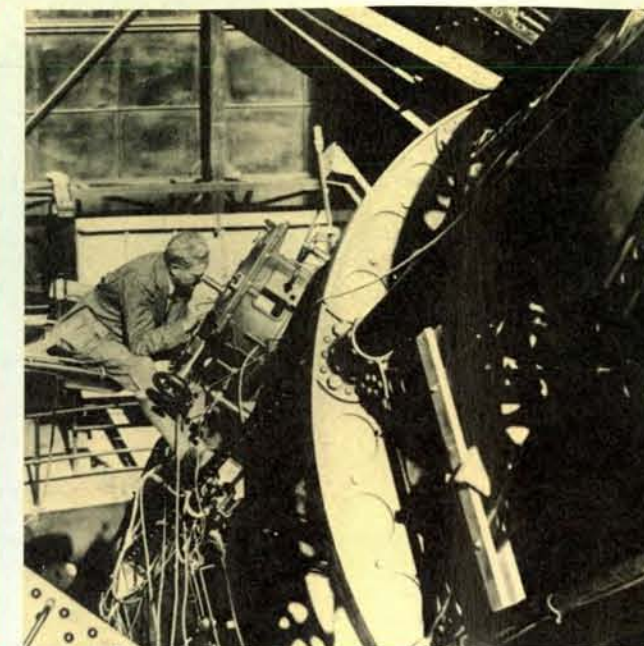
La prima guerra mondiale costrinse Russell a rivolgere le proprie energie all'esame di problemi di navigazione aerea e di fonotelemetria, ma egli cercò un diversivo nel prevedere il futuro dell'astronomia statunitense. Il punto forte delle ricerche condotte in questo paese era il lavoro osservativo; gli Stati Uniti potevano vantare i telescopi più grandi del mondo e i cieli più tersi. L'Harvard College Observatory, il Lick Observatory e lo Yerkes Observatory nel Wisconsin erano intensamente impegnati in censimenti stellari che erano stati avviati prima dell'inizio del secolo.

Russell comprese però che gli Stati Uniti erano in ritardo rispetto all'Europa nell'astrofisica teorica e cominciò a mettere in dubbio i progetti di catalogazione apparentemente inesauribili, che davano una rappresentazione sempre più minuziosa del cielo senza però basi teoriche adeguate. Qualche secolo prima le leggi di Newton avevano permesso di descrivere i moti orbitali dei corpi celesti. Non era possibile che le leggi della fisica delle radiazioni e della meccanica quantistica determinate da Planck, Bohr ed Einstein fossero utilizzabili per descrivere la struttura delle stelle?

Russell e altri pensavano che si dovessero affrontare questi problemi prima di progettare altri programmi osservativi. Nel 1916 egli mise apertamente in discussione le ricerche di tutta una vita del suo vecchio amico e protettore Pickering: «Troppo spesso accade che si comincino lavori di routine prima ancora che i metodi siano stati compiutamente perfezionati». Nel 1917 Hale chiese a Russell di preparare la prima di varie rassegne sullo stato delle ricerche per il National Research Council, che era stato appena fondato. In essa Russell sostenne che molte aree dell'astronomia moderna avrebbero potuto essere rafforzate da connessioni più strette con la fisica teorica. E di nuovo, nel 1917, scrisse a Pickering in merito alle ricerche condotte dalla sua assistente: «Per essere del tutto franco, mi pare che Miss Cannon si sia interessata più all'*aspetto* degli spettri che al loro *significato*. Io non penso che questo fatto sminuisca il servigio da lei reso all'astronomia; al contrario, la sua attenzione rigorosa ai fatti, senza riguardo per le teorie correnti, l'ha resa particolarmente idonea a compiere il suo grande lavoro.»

Pickering, convinto com'era che le esigenze del programma di osservazione di Harvard non fossero state rappresentate in misura adeguata nel National Research Council, ribatté che forse, nell'interesse del progresso dell'astronomia, sarebbe stato bene che «il sognatore suggerisse all'uomo pratico quali fatti desidera». La controversia non si spinse mai oltre questo punto. Quando Pickering morì, nel 1919, Russell gli rese omaggio dicendo che erano state «le informazioni spettroscopiche che [Pickering] mi inviò per le stelle di cui avevo studiato le parallassi - dono disinteressato a uno studente giovane e sconosciuto - a mettermi per la prima volta sulla via che mi avrebbe condotto all'intera teoria delle stelle giganti e nane. Io non credo che si sia mai visto un uomo di scienza più generoso di lui.»

Nonostante questo omaggio a Pickering, Russell rimase saldo nella sua convinzione che l'indirizzo dell'astronomia dovesse mutare. Così aveva scritto a Pickering nel 1917: «Mi pare che l'astronomia attuale sia come un esercito che avanzi con due ali, una lungo la linea dell'osservazione di routine e l'altra lungo quella della ricerca di principi. Se le due ali non si tengono sempre in contatto, l'esercito non andrà lontano.» Russell si assunse il compito di guidare l'ala teorica dell'astronomia statunitense e di trasformarla nella moderna astrofisica.

L'interferometro di Michelson fu montato sulla parte superiore del telescopio di 100 pollici di Mount Wilson nel 1920. La luce proveniente da una stella è riflessa da due specchi mobili posti alle estremità di una trave lunga 6 metri; i due fasci di luce vengono riuniti nel fuoco del telescopio, dove si combinano, ossia interferiscono. Muovendo gli specchi si possono far sparire le frange di interferenza; la separazione degli specchi è allora una misura della dimensione angolare (diametro apparente) della stella. Con questa tecnica il potere risolutivo del telescopio è reso equivalente a quello di un telescopio con uno specchio di 20 piedi di diametro. Questo strumento venne costruito sotto la supervisione di Albert A. Michelson, e alla fine del 1920 Francis G. Pease (qui ritratto) lo usò per determinare per la prima volta il diametro apparente di una stella: Betelgeuse. Gli specchi venivano spostati manualmente da un assistente.

Nonostante la sua retorica, negli anni immediatamente successivi alla prima guerra mondiale, Russell si lasciò trascinare da interessi occasionali. Fra il 1914 e il 1919 pubblicò 26 articoli di astronomia su 15 temi diversi: stelle binarie, l'orbita della Luna, energia ed evoluzione delle stelle, magnitudini e masse stellari, parallassi, stelle variabili e così via. Le sue ricerche erano valide, ma non potevano certo costituire punti fermi nella via verso future acquisizioni.

La relazione del 1917 al National Research Council dimostra che Russell capiva che la chiave per applicare le acquisizioni di Planck, Bohr ed Einstein alla struttura e all'evoluzione delle stelle era l'analisi degli spettri stellari; nondimeno egli si trovava in una situazione irta di ostacoli. A quell'epoca gli astronomi potevano identificare gli elementi chimici che producevano certe righe di assorbimento o di emissione solo confrontando tali righe con quelle ottenute in laboratorio. Essi erano in grado di classificare le stelle in base agli spettri ma solo in modo qualitativo, dall'assenza o dalla presenza di righe e dalla loro intensità. Non si aveva un'idea chiara dei meccanismi fisici che producevano le righe di assorbimento o di emissione. Russell e altri riconobbero che la differenza nelle temperature stellari era la causa prima della disparità degli spettri e che la composizione delle stelle era probabilmente solo una causa secondaria, ma questa concezione incontrava forti resistenze. Non c'erano ancora spiegazioni fisiche del perché mutamenti di temperatura dovessero produrre spettri diversi.

Gli astronomi disponevano di elementi un poco più solidi per quanto concerneva gli spettri continui: il fondo continuo a tutte le lunghezze d'onda al quale si sovrappongono le righe di assorbimento o di emissione di una stella. Soprattutto grazie alle eccellenti ricerche di colorimetria eseguite in Germania, si era

stabilito che la quantità di radiazione emessa da una stella in una data banda di lunghezze d'onda (il suo indice di colore) rassomigliava al famoso spettro del corpo nero di Planck, che è lo spettro di un oggetto teorico le cui proprietà radiative dipendono esclusivamente dalla temperatura. Ma la connessione era ancora debole; inoltre, molte stelle non sembravano comportarsi affatto come corpi neri.

Russell non fu l'unico a riconoscere l'importanza cruciale di una migliore comprensione degli spettri stellari. Anche Arthur Stanley Eddington a Cambridge, in Inghilterra, la pensava allo stesso modo. La formula di Planck dà la quantità di radiazione che un oggetto emette per area unitaria in un angolo solido unitario, ossia la sua luminosità superficiale. Per confermare che una stella si comporta come un corpo nero si deve dunque conoscere l'angolo solido che essa sottende, ovvero il suo diametro apparente. Inversamente, se si suppone che una stella si comporti come un corpo nero e se la sua luminosità superficiale è nota, è possibile ricavarne il diametro apparente.

Nel 1920 Eddington osservò, nel suo discorso da presidente alla British Association for the Advancement of Science, che «probabilmente la cosa più necessaria per l'attuale astronomia stellare, per essere certi che le nostre deduzioni poggino su basi corrette, è un modo per misurare il diametro angolare apparente delle stelle». Egli passò poi a stimare, sulla base dell'indice di colore, la temperatura e la luminosità superficiale di Betelgeuse e mostrò che, se essa si comportava come un corpo nero, doveva sottendere 0,051 secondi d'arco, ossia che il suo diametro apparente doveva essere pari a 1/36 000 del diametro apparente della Luna.

Le conferme vennero molto presto. Tanto Eddington quanto Russell sapevano che a Mount Wilson era in costruzione, sotto la direzione di Albert A. Michelson, uno strumento astronomico radicalmente nuovo, in grado di misurare la separazione di stelle doppie molto strette o addirittura il diametro angolare di una stella: un gigantesco interferometro ottico di 20 piedi destinato a essere applicato al telescopio Hooker di 100 pollici, inaugurato da poco (*si veda l'illustrazione a pagina 89*). Hale, che si trovava allora a Pasadena, aveva appena posato la sua copia del 2 settembre di «Nature», dove era stampato l'indirizzo di Eddington, quando prese la penna per informare Michelson che finalmente c'era una stima teorica accettabile per il diametro apparente di una stella. In dicembre l'interferometro di Michelson entrò in attività e la stima di Eddington fu confermata. Questa fu la prima verifica sperimentale del fatto che le stelle si comportano effettivamente secondo le leggi della fisica moderna.

Anche Russell aveva fatto una stima teorica del diametro di Betelgeuse, che fu pubblicata verso la fine di dicembre. Essa era vicina al valore osservato, ma quella di Eddington risultò più esatta. Benché Michelson avesse in seguito riconosciuto il valore della stima di Russell, questi si rese conto da quest'episodio che il futuro dell'astronomia sarebbe stato legato in gran parte ai metodi della scuola di Cambridge, che era all'avanguardia nell'applicazione della fisica allo studio delle stelle.

Nel dicembre 1920 vi fu un altro evento fondamentale. A Russell capitò fra le mani una copia del numero di ottobre del «Philosophical Magazine», dove uno sconosciuto giovane fisico indiano di Calcutta, Meghnad N. Saha, associava audacemente il potenziale di ionizzazione di un elemento chimico (l'energia necessaria per staccare un elettrone da un atomo) e il suo grado di ionizzazione alla temperatura e alla pressione dell'ambiente circostante. Russell riconobbe subito che questa era la chiave che egli andava cercando per la comprensione degli spettri stellari. Egli scrisse allora a Walter S. Adams, uno dei suoi collaboratori a Mount Wilson: «Credo che in capo a qualche anno potremo utilizzare la conoscenza dei potenziali di ionizzazioni, e via dicendo, per ottenere determinazioni numeriche di temperature stellari a partire da dati spettroscopici».

Lavorando nel quadro teorico dell'atomo di Bohr, Saha aveva applicato alle atmosfere stellari i concetti di equilibrio termodinamico e di ionizzazione termica. Secondo il modello di Bohr, un fotone che urti un atomo può eccitare un elettrone facendolo passare a un livello di energia più elevato; nel corso di questo processo il fotone viene assorbito. Inversamente, quando un elettrone scende da un livello di energia superiore a uno inferiore viene emesso un fotone. Saha capì che questi concetti potevano aiutare a far luce sul comportamento degli atomi nelle atmosfere stellari. In particolare, mostrò perché le righe spettrali di un dato elemento siano più intense a certe temperature e più deboli ad altre.

Per esempio, in un'atmosfera stellare molto molto più fredda di 4000 kelvin, la maggior parte degli atomi di idrogeno ha gli elettroni nello stato energetico più basso. (Queste stelle rientrano nelle classi K e M nella familiare sequenza O, B, A, F, G, K, M, stabilita ad Harvard, dove le stelle O sono le più blu e le M sono le più rosse.) In queste stelle relativamente fredde i fotoni non sono in grado di eccitare transizioni fra i livelli più elevati e di essere assorbiti nel processo; perciò sono assenti le righe di assorbimento corrispondenti alle transizioni fra i livelli superiori. A 10 000 kelvin circa si verificano transizioni fra i livelli superiori e lo spettro visibile è dominato da righe di assorbimento (la cosiddetta «serie di Balmer»). In stelle molto più calde (della classe A) l'idrogeno è in gran parte ionizzato - o esiste in alti stati di eccitazione - ogni transizione si verifica al di sopra delle frequenze visibili e le righe di assorbimento dell'idrogeno tornano a indebolirsi. L'alta pressione tende a ridurre la quantità di ionizzazione, cosicché la pressione nella stella influisce anche sulle righe di assorbimento e di emissione.

All'articolo di ottobre sulla ionizzazione termica Saha ne fece seguire in rapida successione altri due mentre era ancora a Calcutta. Poi si recò in Inghilterra, sperando di lavorare a Cambridge, ma fu aiutato solo da Alfred Fowler dell'Imperial College of Science and Technology di Londra. Un quarto articolo uscì dopo che Saha era entrato a far parte del laboratorio di Fowler. In esso si stabiliva, come è implicito in quanto detto sopra, che la classificazione O, B, A, F, G, K, M rappresenta non solo una sequenza di colore dal blu al rosso, ma anche una sequenza di temperature assolute, con le stelle O all'estremo più caldo e le M a quello più freddo.

Mentre si trovava a Londra, Saha ebbe a lamentare: «Praticamente non abbiamo dati di laboratorio che ci guidino». Anche Russell si era accorto della mancanza di dati di alta qualità sugli spettri e sulla ionizzazione. A differenza di Saha, però, sapeva dove andare a procurarseli: da George Ellery Hale e dai suoi astronomi a Mount Wilson.

Già nella sua lettera di dicembre ad Adams, Russell aveva identificato un modo per utilizzare le scoperte di Saha. Adams supponeva da molto tempo che differenze di pressione nelle atmosfere stellari potessero influire sull'aspetto degli spettri. Il suo metodo delle parallassi spettroscopiche, da lui sviluppato con Anton Kohlshutter e in cui certe righe spettrali potevano servire come indicatori di distanze, lasciava intendere il ruolo della pressione negli spettri stellari; esso veniva ora posto su una solida base teorica dalla proposta di Saha. Russell sottolineò questa conclusione e rimandò Adams alla sua grande collezione di spettri stellari a Mount Wilson, suggerendogli di cercare l'idrogeno molecolare in stelle molto fredde.

L'estate seguente lo stesso Russell si recò a Mount Wilson come ricercatore associato per applicare alle stelle la teoria di Saha. Nei successivi due decenni Russell lasciò Princeton almeno una volta all'anno, talvolta due, per recarsi agli uffici del Mount Wilson Observatory a Pasadena. Nel corso dei suoi viaggi di trasferimento si fermava spesso in altri osservatori per dare consigli ai ricercatori, per andare alla ricerca di dati e per tenere conferenze su una grande varietà di argomenti. Pur conservando alcuni dei suoi precedenti interessi di ricerca, come quello per le stelle binarie e variabili, egli si pose il nuovo obiettivo di sviluppare una teoria degli spettri stellari. Nella prima estate da lui trascorsa a Mount Wilson, disse in una conferenza a Berkeley che, con la teoria di Saha, «l'astronomia, la fisica e la chimica hanno ora un modello atomico per l'emissione e l'assorbimento di radiazione». Russell aveva già usato la collezione di immagini spettroscopiche del Sole a Mount Wilson per dimostrare la correttezza di alcune previsioni di Saha e per affermare che in generale la teoria poteva spiegare il comportamento degli elementi nell'atmosfera solare. Egli aggiunse: «Questo è solo un esempio delle immense possibilità del nuovo campo di ricerca che si dischiude dinanzi a noi. Saranno necessari molti studi prima che esso possa essere esplorato anche sommariamente, e l'astronomo, il fisico e il chimico dovranno unirsi in questo attacco... Non è troppo audace sperare che, nell'arco di pochi anni, la scienza possa

12 SCIENTIFIC AMERICAN July 2, 1921

The Heavens for July, 1921

What a Study of Atoms and Electrons Tells Us of the Stars

By Henry Norris Russell, Ph.D.

IT is becoming more and more evident, as both sciences advance, that the astronomy of the future will be intimately associated with and dependent upon the concepts and the results of physics, and especially of that branch of physics which deals with the constitution and properties of atoms. Our knowledge within the latter field has been very greatly extended within the last decade, and many astronomical observations which before were puzzling have thereby been explained.

This is particularly true in the realm of spectroscopy. The main facts regarding the emission of light by hot bodies, and by hot gases in particular, have been known for many years; but it is only recently that we have even begun to have an idea of the processes taking place inside the atoms of the gas, which are involved.

For example, when the vapor of a given element, such as calcium or iron, is confined in a heated tube or "furnace" and observed through the end of the tube, the spectrum of the light which it emits shows certain bright lines. If the temperature is raised these lines grow stronger and new lines appear in addition. When the same metal is brought into an electric arc (which is hotter, and also subject to direct electrical action), more lines appear; while a yet more advanced stage may be reached by passing a powerful spark, fed by a source of current of high tension, between two bits of the metal; and in the spectrum from this lines may be found which were not to be observed at any of the lower stages of temperature.

Extensive studies have been made of these phenomena, and long lists of "furnace" and "spark" lines compiled, with important astronomical applications. But the physical explanation, from the atomic standpoint, lagged behind, and came only with the application of the modern quantum theory, which has been remarkably successful.

Why Are the Spectral Lines?

We have good reason to believe that an atom of any element consists of a central, and very small, nucleus, carrying a positive electrical charge, surrounded by a number of negatively charged electrons, which under the system of forces acting between them and the nucleus arrange themselves automatically in a definite pattern, probably consisting of several concentric shells or layers, at least in the heavier atoms. In the hydrogen atom there is but one electron; in helium two; in oxygen eight; in sodium eleven; in iron twenty-six; and so on up to 82 for lead and 92 for uranium. The inner electrons are held by very powerful forces, and are hard to dislodge; but a few of the outermost are relatively easy to displace, and it is these which are concerned in the chemical affinity between atoms of different sorts, and also in the production of the radiation of the visible spectrum. To pull one of these electrons away from the rest of the atom, or as it is called to ionize the atom, demands a certain expenditure of energy; and this produces an absorption of light by the gas of which this atom is a part. When some other free electron comes near to the ionized atom, it will be attracted to it (provided it does not go by too fast); and, in falling back, a corresponding amount of energy will be emitted in the form of light radiated by the gas.

Recent research has shown that this is but part of the story. There appear to be many different positions in which the electron can stop, short of being pulled clear away from the atom. The farther out it gets the more energy is required to raise it—the greatest amount of all corresponding to the complete removal of the electron, or the ionization of the atom.

Now when an electron changes from one of these states to another, light is absorbed, if it is pulled up to a "higher level" nearer the outside of the atom, or emitted if it drops to a "lower level"; and this light consists of vibrations at a perfectly definite rate, giving a sharp line in the spectrum. The most remarkable feature remains to be mentioned. The number of light vibrations per second is exactly proportional to the amount of energy which is required to pull the electron up from one position to the other, or is liberated when it comes back. The reason for this famous "quantum relation"—and indeed the reason why the various possible positions for the electron should exist at all—remains still a mystery, which is regarded by the ablest physicists as one of the hardest problems of science. But the fact has been tested in so many ways that no doubt remains.

When the spectra of the elements are studied from this standpoint it is found that the furnace lines correspond (in the case of absorption) to the raising of the electron from the very lowest "level" at which it normally is situated in the undisturbed atom to various higher levels; while the arc lines, in general, correspond to the raising of the electron from one of these higher levels to another. When light is emitted we have to do with an electron falling back over one of the same intervals.

The enhanced lines correspond to still another process. After one electron has been taken clear out of the atom, it is often possible, by a greater force, to pull a second electron out, and doubly ionize the atom. In this process too there are various possible "levels" between which the second electron may shift, and a corresponding set of lines, all quite different from the furnace or arc lines. It is even possible that an atom may lose a third or actually a fourth electron, and there is reason to suppose that some spectral lines, produced only in very violent sparks, are of this origin.

At 11 o'clock: July 7.
At 10½ o'clock: July 14.
At 10 o'clock: July 22.
At 9½ o'clock: July 30.
At 9 o'clock: Aug. 7.
At 8½ o'clock: Aug. 14.
At 8 o'clock: Aug. 22.
The hours given are in Standard Time. When local summer time is in effect, they must be made one hour later: 12 o'clock on July 7, etc.
NIGHT SKY: JULY AND AUGUST

What It Means to the Astronomer

With these ideas in mind it is very easy to see why the furnace lines are characteristic of the red stars, like Betelguese; the arc lines of yellow stars, like the sun; and the spark lines of very white stars, like Sirius. In the hot atmospheres of the stars, the atoms collide and jostle one another. The red stars are the coolest, and the collisions are the least violent, so that most of the atoms are in their undisturbed condition, and absorb only the flame lines. In the hotter atmosphere of the sun many of the atoms are jostled so that the electrons within them are raised to higher "levels" and are in a position to be raised further, with absorption of the arc lines. Finally, in the still hotter white stars most of the atoms are completely ionized, and are therefore ready to have a second electron removed, with absorption of the light corresponding to the spark lines. For some elements, such as calcium, this process occurs with relative ease; hence the spark lines of calcium—the great H and K lines in the violet—appear strongly in the sun. Helium on the other hand is the most difficult of all the elements to ionize; and the amount of energy required even to lift an electron from the lowest "level" to the next above is so great that the corresponding light vibrations are exceedingly rapid, and lie so far in the ultra-violet that all ordinarily transparent substances are opaque for them. The visible lines of helium correspond to a lifting of an electron from the second, or even a higher level to one still above, and can only be produced in an atom which has already been violently jostled, so as to throw the electron up to the second "level." This explains why the absorption lines of helium are found only in the very hot stars, like those in Orion. Spark lines of helium, corresponding to the loss of a second electron, are known; but these are found only in a very few stars which, from other evidence as well, we have reason to believe to be the hottest in the heavens.

Many beautiful applications of this theory have recently been worked out by an Indian physicist, Dr. Megh Nad Saha, of the University of Calcutta. Much of the foregoing discussion is adapted from his work, and one more instance of it may be given. The dark lines of sodium are strong in the solar spectrum. Those of potassium are present, but weak. The rare alkali metals, rubidium and caesium, show many strong lines but these do not appear in the sun at all. This has long been a puzzle, but Dr. Saha has given the solution.

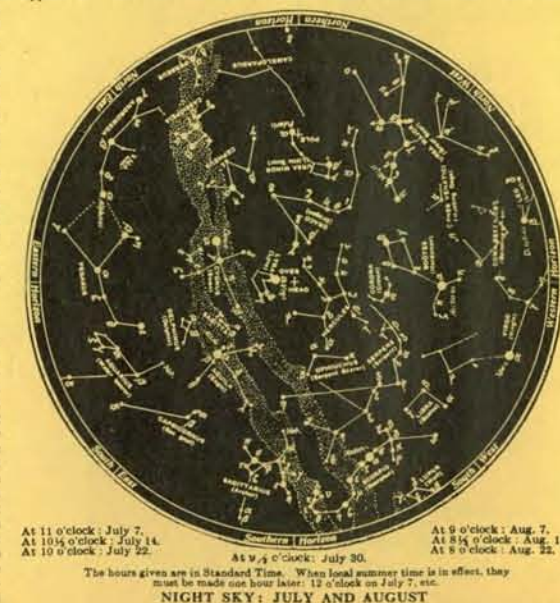
Laboratory experiments have shown that it is fairly easy to remove an electron from a sodium atom, easier to get one out of a potassium atom, and still easier for rubidium and caesium. To get a second electron away from any of these atoms, after the first is gone, is however very difficult. Calculation shows that, in the sun's atmosphere, sodium vapor should be largely ionized, with however a considerable percentage remaining un-ionized atoms, which still retain one electron that may be removed by the action of light, with absorption of the well-known sodium lines. For potassium, almost all the atoms are ionized, leaving very few in a position to produce the absorption lines. Rubidium and caesium, still easier to ionize, would be completely ionized, leaving no atoms at all in a position to produce the absorption lines which are so conspicuous under the less extreme conditions of our laboratories. Hence the weakness of the potassium lines, and the absence of those of the other elements, is completely explained.

When more laboratory work has been done (largely by electrical methods) on these matters, it probably will be possible to calculate with fair precision the temperatures of the atmospheres of the various types of stars, simply from a knowledge of the degree to which the various sorts of atoms in them are ionized, as indicated by the lines in their spectra.

The Heavens

At our hour of observation Vega is almost overhead. Cygnus is high in the east, and Aquila in the southeast, a little lower. Below it lie Capricornus and Aquarius, and to the right, due south, is Sagittarius, with Scorpio to the west of it, and Ophiuchus above the latter. Bootes is the most conspicuous western constellation, with Corona above it and Hercules almost overhead. Ursa Major is in the northwest, Ursa Minor and Draco in the north, Cassiopeia and Cepheus in the northeast, and Pegasus has just risen in the east.

The Planets

Mercury is an evening star at the beginning of the
(*Continued on page 16*)

MEGHNAD N. SAHA, quinto figlio di un negoziante, nacque nel 1893 nell'attuale Bangladesh. Avviato agli studi universitari, manifestò un grande talento nelle discipline da lui scelte della matematica e della fisica; ebbe un temperamento aggressivo e spesso si trovò in contrasto con i superiori a causa delle sue idee politiche e di conflitti di personalità. Laureatosi nel 1918 all'University of Science di Calcutta, ben presto si volse allo studio di problemi di astrofisica, in cui apportò ciò che aveva appreso tenendo corsi di perfezionamento in termodinamica e meccanica quantistica. In una serie di articoli scritti fra il 1919 e il 1921 a Calcutta e in Inghilterra, applicò il concetto di ionizzazione termica alle atmosfere stellari e preparò la via alla comprensione degli spettri stellari; questi articoli vengono considerati talvolta l'inizio dell'astrofisica moderna.

Nonostante l'importanza dei suoi lavori, Saha non fu invitato dai colleghi europei e americani a collaborare con loro nel perfezionamento e nello sviluppo della teoria. Tornato in India e all'Università di Calcutta, poco tempo dopo divenne professore di fisica all'Università di Allahabad. Nei due decenni successivi divenne una figura centrale nella creazione del National Institute of Sciences of India, della Indian Physical Society e della Indian Science News Organization. Dopo essere tornato all'Università di Calcutta nel 1938, fondò un istituto per lo studio della fisica nucleare e si batté a sostegno della necessità di una pianificazione nazionale su una base scientifica. Ebbe vasta influenza in aree che andavano dal controllo delle inondazioni e dalla pianificazione economica alla riforma del calendario indiano e agli usi pacifici dell'energia atomica. Nel 1951 fu eletto al Parlamento dell'India come deputato indipendente e vi rimase sino alla morte, avvenuta nel 1956.

giungere in possesso di una teoria razionale degli spettri stellari e, al tempo stesso, di molte conoscenze in più sulla costituzione degli atomi.»

Russell non fu l'unico a scorgere queste possibilità. Lo stesso Saha sollecitò aiuti per poter proseguire le proprie ricerche e scrisse a Hale per chiedere le stesse cose che erano state fornite a Russell. Al tempo stesso anche Ralph H. Fowler e Arthur Milne in Inghilterra riconobbero le potenzialità dell'opera di Saha e negli anni seguenti lavorarono per completare la teoria. Essi notarono che la formula di Saha che descriveva le classi spettrali come una sequenza di temperature assolute non teneva conto in modo appropriato del fatto che in un'atmosfera stellare era presente più di un elemento e rimediarono a questa omissione. Al tempo stesso perfezionarono la parte della teoria concernente il ruolo della pressione.

Molti altri fondarono le proprie ricerche sui lavori di Saha, ma fu a Mount Wilson, con la coordinazione di Russell e degli astronomi di Hale, che ebbe luogo l'«attacco generale agli spettri». Per fare un favore a Hale, Russell rispose alla lettera di Saha, spiegandogli per sommi capi quale fosse il piano di lavoro per Mount Wilson. Gli assicurò che avrebbero seguito la sua guida, ma Saha non fu invitato a partecipare alla ricerca.

Poco tempo dopo aver verificato le previsioni di Saha, Russell seguì le orme di Fowler e Milne nel tentativo di perfezionare la teoria stessa e anche di spiegare le anomalie spettroscopiche lasciate in sospeso nella versione originale di Saha. Si presentarono subito due rompicapo: il comportamento inatteso del bario e la persistenza dell'idrogeno nello spettro di tutte le stelle.

Russell trovò che il bario era più ionizzato del sodio nello spettro solare, cosa strana perché i due elementi hanno lo stesso potenziale di ionizzazione. Il rompicapo del bario condusse Russell a occuparsi di fisica teorica per classificare gli spettri degli elementi alcalino-terrosi, che comprendono il bario.

Gli elementi alcalino-terrosi si distinguono per il fatto di avere due elettroni di valenza, o elettroni esterni, anziché uno. In collaborazione con F. A. Saunders, Russell derivò un modello migliorato della struttura di atomi in cui due elettroni anziché uno partecipavano alla produzione di righe spettrali. Le regole per l'interazione di due elettroni escogitate da Russell nel corso di questa collaborazione sono chiamate «accoppiamento di Russell-Saunders»; esse permisero di giustificare gli spettri del bario e degli elementi alcalino-terrosi. Ancor prima di completare questa ricerca, Russell fece il passo successivo: quello di esaminare gli spettri prodotti da atomi con tre elettroni di valenza. Egli scelse il titanio e trovò un'altra regola di transizione. Russell era entusiasta; non solo le sue teorie fisiche spiegavano come funzionassero le stelle, ma gli spettri stellari potevano anche essere usati come strumenti per sondare i misteri dell'atomo.

Mentre, per quanto riguardava bario e titanio, si poté parlare di grandi successi di Russell, la questione dell'idrogeno rimaneva una spina nel suo fianco. A quell'epoca gli astronomi ritenevano che nessun elemento dominasse nelle atmosfere stellari, che si pensava fossero miscele gassose di elementi generalmente pesanti, in particolare ferro. Inoltre, la teoria della struttura stellare di Eddington richiedeva che il peso molecolare medio del gas dovesse essere molto maggiore di quello dell'idrogeno. Eppure l'idrogeno persisteva virtualmente in tutti gli spettri stellari.

Il rompicapo dell'idrogeno condusse Russell e gli astronomi di Mount Wilson a ricalibrare le lunghezze d'onda dello spettro del Sole sugli standard di laboratorio. A Princeton, Russell e la sua infaticabile assistente Charlotte E. Moore calibrarono l'intensità delle righe spettrali del Sole in base a una nuova teoria, sviluppata da Russell e altri, che spiegava l'intensità delle righe in funzione delle concentrazioni relative degli elementi presenti nell'atmosfera solare.

Ma il rompicapo dell'idrogeno non era risolto. Nonostante la prova evidente delle righe spettrali, Russell nutriva dubbi sul fatto che l'idrogeno predominasse nell'atmosfera di tutte le stelle. Mentre si sforzava di venire a capo del problema dell'idrogeno, Russell aveva mandato Menzel a Harvard per sfruttare la collezione incomparabile di spettri stellari di quell'osservatorio. Al tempo stesso era arrivata da Cambridge una giovane astronoma, Cecilia Payne, mandata da Eddington, la quale portava con sé le nuove idee di Fowler e di Milne. Anch'ella aveva in animo di esaminare la struttura dell'atomo con l'aiuto degli spettri conservati ad Harvard e di determinare gli elementi presenti nelle stelle meglio di quanto fosse riuscito a fare Saha. La Payne riuscì a realizzare tutti i suoi propositi nella sua monumentale tesi di dottorato del 1925; fu lei a riconoscere per la prima volta che l'idrogeno è di gran lunga più abbondante di ogni altro elemento nell'atmosfera delle stelle.

La sua conclusione non era facile da accettare. Essa minacciava la teoria della struttura stellare di Eddington, cosa che indusse Russell a dirle che il suo risultato era chiaramente impossibile. La Payne seguì disciplinatamente i consigli di Russell nella pubblicazione della sua tesi, ma in privato rimase fedele alle proprie conclusioni.

Il problema dell'abbondanza dell'idrogeno tornò a proporsi varie volte e a frustrare gli sforzi di Russell, che tentò di ignorarlo per tutto il 1928, anche se molti suoi colleghi avevano già cominciato a credere che le conclusioni originarie della Payne fossero corrette. Russell mise in campo tutte le sue forze a Princeton e a Mount Wilson per condurre il suo attacco all'idrogeno. Infine, in un importantissimo articolo del 1929, che egli definì come «una ricognizione in un territorio inesplorato», raccolse tutte le prove spettroscopiche accumulate da lui e dalla Moore e dichiarò che, in definitiva, le atmosfere stellari erano composte principalmente da idrogeno.

Durante il suo attacco finale all'idrogeno, Russell fu consapevole del fatto che Albrecht Unsöld - un allievo del fisico teorico Arnold Sommerfeld - era in grado di derivare informazioni sulle abbondanze assolute degli elementi e dei composti dai profili delle righe spettrali, cosa che Russell non aveva mai tentato di fare. Anche Unsöld stava confermando che le atmosfere stellari sono composte prevalentemente da idrogeno e Russell sapeva che le tecniche del giovane tedesco erano molto più efficaci delle sue.

La piena padronanza della nuova fisica quantistica che Unsöld poteva vantare rappresentava il futuro degli studi sulle atmosfere stellari. Unsöld e vari altri scienziati europei avrebbero infine perfezionato la prima rozza stima di Russell delle abbondanze relative nell'atmosfera solare. Russell era molto compiaciuto di queste estensioni della sua opera da parte di altri, ma era evidente che la nuova «meccanica quantistica» stava prendendo rapidamente il sopravvento. Alla metà degli anni venti, mentre Russell stava cercando una teoria degli spettri di multipletto (spettri prodotti da livelli di energia atomica fittamente addensati), Sommerfeld, H. Hönel e R. de L. Kronig erano già sulla strada giusta e riuscirono a pubblicare i loro risultati prima di lui.

Tali competizioni teoriche diedero a Russell la consapevolezza della schiera di fisici europei che stavano affrontando il problema della struttura dell'atomo. Russell sapeva di essere sovverchiato dal numero e in buona misura distanziato. Questa situazione era in parte una conseguenza della sua avversione per la direzione in cui la fisica si stava incamminando. Russell non si sentì mai a suo agio con i complessi formalismi matematici della meccanica quantistica ed ebbe sempre una maggiore simpatia per quello che chiamava il modello astronomico dell'atomo di Bohr, fino alla sua metafora dello «spin» (l'equivalente della rotazione). Molto tempo dopo Russell, seguendo Max Born, si riferiva ancora al principio di indeterminazione di Heisenberg come a un «principio di misurabilità limitata». E neppure si sentì mai perfettamente a suo agio con il dualismo onda-particella della materia, anche se era disposto ad applicare entrambi i modelli a «problemi pratici».

Dopo la metà degli anni venti, Russell preferì lasciare ad altri il compito di apportare sviluppi alla teoria. Egli continuò ad ammirare l'efficacia e la generalità del modello di Bohr e la sua capacità di fornire regole per il calcolo degli spettri atomici. Come tanti spettroscopisti della sua generazione, Russell si accontentò di risolvere rompicapo, applicando il modello di Bohr agli spettri nel tentativo di svelare la struttura degli atomi.

Il ruolo di Russell come pioniere dell'astrofisica quantitativa fu quello di chi indica la via da seguire. Egli fu una figura di transizione, ma non compì mai tale transizione personalmente; alcuni suoi allievi riuscirono però a compiere questo passo, e gran parte della durevole influenza di Russell si esercitò attraverso di loro: astronomi come Spitzer e Menzel e i loro giovani allievi.

I molti ruoli ricoperti da Russell indussero Shapley a conferirgli il titolo di «decano degli astronomi statunitensi». In Russell c'era un insolito connubio di due stili scientifici classici: la profondità e l'ampiezza di vedute. Egli guardava con simpatia alla necessità di osservazioni programmatiche, ma diffidava di vasti progetti non ispirati a teorie. Russell ebbe spesso una lucidità di visione che gli permise di suggerire ad altri linee di ricerca feconde. Come ebbe a dire una volta Cecilia Payne-Gaposchkin in età matura: «Henry Norris Russell sapeva riconoscere una buona occasione quando la vedeva». La capacità di trarre vantaggio da queste «buone occasioni» e di saper comunicare ad altri il bisogno di una ricerca sistematica a lungo termine ispirata da una teoria fisica permise a Russell di esercitare una vasta influenza. In una rassegna di astronomi pubblicata nel 1946, Russell era citato per lo più come un insegnante particolarmente stimolante; ancora negli anni sessanta e settanta, molto tempo dopo la sua morte, avvenuta a 79 anni nel 1957, la sua opera continuò a essere citata a una media di 50 volte all'anno.

BIBLIOGRAFIA

MENZEL DONALD H., *The History of Astronomical Spectroscopy* in «Annals of the New York Academy of Sciences», 198, pp. 225-244, 1972.

HUFBAUER KARL, *Astronomers Take up the Stellar Energy Problem* in «Historical Studies in the Physical Sciences», 11, parte 2, pp. 277-303, marzo 1981.

DEVORKIN DAVID H. e KENAT RALPH, *Quantum Physics and the Stars* in «Journal for the History of Astronomy», 14, n. 40, parte 2, pp. 102-132, giugno 1983 e 14, n. 41, parte 3, pp. 180-222, ottobre 1983.

DEVORKIN DAVID H., *Stellar Evolution and the Origins of the Hertzsprung-Russell Diagram in Early Astrophysics* in *Astrophysics and Twentieth-Century Astronomy to 1950* a cura di Owen Gingerich, Cambridge University Press, 1984.

(RI)CREAZIONI AL CALCOLATORE

di A. K. Dewdney

Evoluzione simulata: microbi che imparano a cacciare batteri

«Per chi, come me, non è un matematico, il calcolatore può costituire un efficace aiuto all'immaginazione.»
RICHARD DAWKINS,
L'orologiaio cieco

Sul fondo fangoso di una pozza d'acqua stagnante si muovono numerosissimi protozoi, cibandosi dei batteri che lentamente cadono su di loro. I protozoi si assomigliano tutti, ma il loro comportamento mostra notevoli differenze. Alcuni si muovono in modo irregolare alla ricerca di batteri, e di conseguenza mangiano poco. Altri si spostano in modo più mirato, seguendo uno schema di ricerca che appare quasi metodico, e riescono a nutrirsi in abbondanza. I mondi microscopici di questo genere hanno un loro fascino, ma questa particolare scena ha un significato speciale: i protozoi metodici si sono evoluti dai loro consimili erratici nello spazio di una sola ora!

Come forse qualche lettore ha già capito, questa scena non è osservata attraverso un microscopio bensì sullo schermo di un calcolatore: viene generata infatti da un programma chiamato Evoluzione simulata, scritto da Michael Palmiter, un insegnante di scuola superiore di Temple City in California. Piccoli protozoi bianchi, che Palmiter chiama microbi, nuotano nello schermo ingoiando batteri viola. Con il succedersi delle generazioni di microbi, è possibile osservare l'evoluzione di nuovi comportamenti alimentari.

Anche Richard Dawkins, dell'Università di Oxford, ha cercato di capire meglio l'evoluzione mediante programmi che tentano di simulare i suoi diversi aspetti. Uno di questi programmi, scritto dallo stesso Dawkins, è stato argomento delle nostre «(Ri)creazioni al calcolatore» più di un anno fa (si veda «Le Scienze» n. 236, aprile 1988). Il programma di Dawkins visualizza biomorfi, ossia forme generate dal calcolatore che a volte somigliano a creature viventi. La loro evoluzione avviene attraverso un processo di «selezione artificiale»: l'operatore sceglie arbitrariamente, come base per le future generazioni di biomorfi, una tra le nove possibili forme varianti del biomorfo presente a un dato momento sullo schermo.

I biomorfi che emergono dal programma di Dawkins possono essere bizzarri e divertenti - e a volte anche simili a vere forme viventi - ma non si può dire che la loro evoluzione sia avvenuta in modo naturale, cioè sulla spinta di pressioni selettive interne. Dawkins, però, è convinto che sia possibile scrivere un programma per calcolatore che imiti la selezione naturale. Le specie generate dal calcolatore e dotate di questa «capacità evolutiva» darebbero luogo a forme sempre più complesse, che la selezione ridurrebbe a un numero ragionevole. Inoltre, i discendenti sopravvissuti dovrebbero avere la capacità di evolversi in modi nuovi, assolutamente impossibili per i loro antenati.

I microrganismi di Palmiter ci fanno certamente fare un passo nella direzione proposta da Dawkins. Guardando l'illustrazione della pagina a fronte ci si accorge che i microbi (i puntini bianchi) vivono in un rettangolo su cui continuano a depositarsi dei batteri (i puntini viola). I microbi vivono muovendosi incessantemente e cibandosi di batteri. Ciascun batterio mangiato da un microbo fornisce a quest'ultimo 40 unità di energia, sufficienti a fargli compiere 40 movimenti. Là dove il cibo è abbondante, un microbo può facilmente acquisire 1500 unità di energia in pochi minuti, ma a quel punto scatta uno strano meccanismo: se il microbo mangia ancora, non trae più alcun beneficio dal nutrimento, finché il suo livello di energia non scende al di sotto delle 1500 unità.

D'altra parte, può capitare che un microbo trovi pochissimo da mangiare per un lungo periodo. In questo caso le sue riserve di energia possono scendere gradualmente a zero. Arrivato a quel punto, esso rimarrebbe malinconicamente immobile per qualche ciclo, quasi a meditare sulla propria fine, per poi svanire definitivamente con un lampo, come una piccola lampadina.

Il successo di un microbo nella ricerca di cibo dipende, naturalmente, dall'abbondanza relativa di batteri nelle sue immediate vicinanze. Dato che i batteri si depositano più o meno uniformemente nel rettangolo, se ne vengono mangiati molti in una certa zona essi saranno necessariamente presenti in gran numero altrove. Alcuni microbi raggiungono le aree relativamente più ricche in minor tempo di altri. Tutto dipende dai movimenti compiuti dal microbo, dal suo schema di ricerca, per così dire.

Lo scenario darwiniano del programma Evoluzione simulata, anche se in modo astratto, è centrato sui «geni» che regolano il modo in cui un microbo si muove. È probabile che questi particolari geni non esistano nei veri protozoi, ma i microbi di Palmiter ne hanno sei, contrassegnati con *A, D, DS, I, SS* e *S* per Avanti, Destra, Destra Stretta, Indietro, Sinistra Stretta e Sinistra. (Tutte le direzioni sono espresse dal punto di vista del microbo. Le curve normali sono di 60 gradi nell'una o nell'altra direzione, mentre le curve strette sono di 120 gradi.)

A ogni data mossa, il microbo punta in una direzione scelta a caso: il programma estrae a sorte una delle sei possibili direzioni da una specie di cappello

Le curve compiute da un microbo sono relative alla sua direzione in un dato istante

a cilindro matematico. Se il programma sceglie *S*, per esempio, il microbo gira a sinistra di 60 gradi. La probabilità che sia scelta una certa direzione è data da un valore assegnato al gene corrispondente. Ne consegue che più alto è il valore di un gene, maggiore è il suo contributo allo schema complessivo di movimento del microbo. Se un microbo, per esempio, ha un valore di *S* elevato rispetto ai valori degli altri cinque geni, le sue svolte a sinistra saranno particolarmente numerose.

Ogni possibile combinazione di valori dei geni dà luogo a un diverso schema generale di movimento, e questa costituzione genetica, qualunque essa sia, caratterizza il microbo per tutta la vita. Esso può solo sperare (se vogliamo dargli un tocco di antropomorfismo) che le cose migliorino per la sua prole.

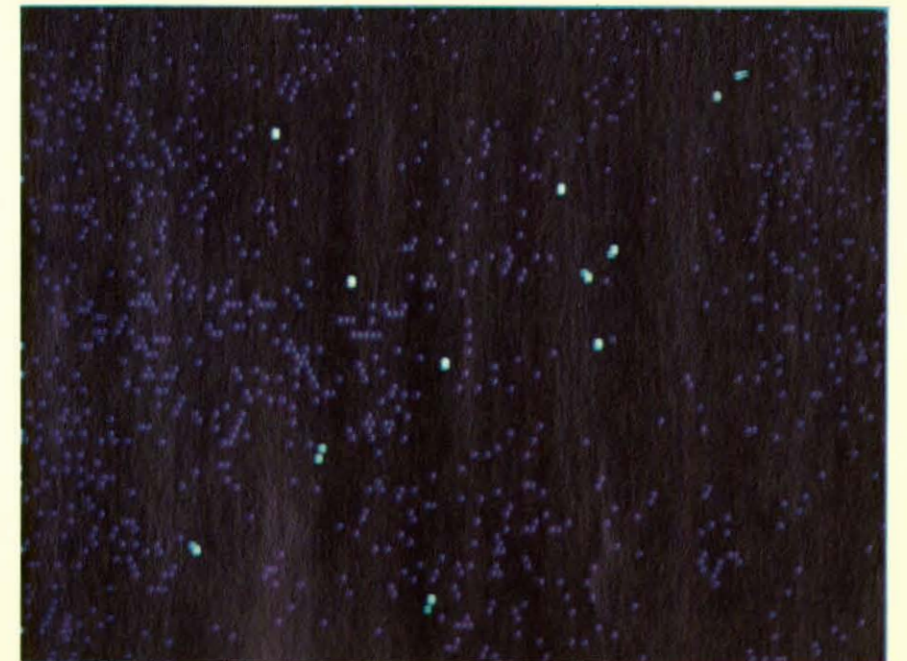
Dopo aver compiuto 800 movimenti, un microbo diventa «maturo» ed è pronto per riprodursi. Per poterlo fare deve anche essere «forte», cioè deve avere almeno 1000 unità di energia immagazzinate sotto la sua membrana di color bianco elettrico. I parameci si riproducono con un processo detto di coniugazione; i microbi, invece, subiscono una scissione: un microbo forte e maturo si divide in due nuovi microbi, ciascuno dei quali possiede metà dell'energia del genitore. Quando si verifica la scissione, i nuovi microbi ereditano dal genitore i geni che ne controllano il movimento, ma con una piccola differenza: c'è un leggero aumento o una leggera diminuzione nel valore di uno dei geni di ciascuno dei due nuovi nati.

Supponiamo, per esempio, che un microbo forte e maturo abbia i seguenti valori genici: $A = 3$, $D = 2$, $DS = 0$, $I = -2$, $SS = 0$ e $S = 1$. I due figli, indicati con *X* e *Y*, potrebbero ereditare le seguenti forme modificate di quella costituzione genetica:

X: $A = 4$ $D = 2$ $DS = 0$ $I = -2$ $SS = 0$ $S = 1$
Y: $A = 3$ $D = 2$ $DS = 0$ $I = -2$ $SS = -1$ $S = 1$

Come si può vedere, in *X* il valore *A* è stato aumentato di 1 e in *Y* è stato diminuito di 1 il valore *SS*.

In che cosa differiranno i figli dal loro genitore? *X* avrà una tendenza leggermente maggiore a spostarsi in avanti, mentre *Y* avrà una tendenza leggermente inferiore a fare curve strette verso sinistra. Queste differenze sono così piccole che anche a un osservatore allenato risultano a mala pena percepibili sullo schermo del calcolatore.

Nella sua forma più semplice, Evoluzione simulata inizia fornendo a 10 microbi una struttura genetica casuale che, nella maggior parte dei casi, li fa vagare da una zona all'altra in modo imprevedibile. Di regola, questi «microbi vaganti» hanno un alto tasso di mortalità: la loro tendenza è quella di mangiare la

I «vaganti» si evolvono lentamente in «incrociatori»

maggior parte del cibo nelle loro immediate vicinanze, per poi ondeggiare fino all'inedia su un substrato deserto. Ciononostante, ce ne sono alcuni che riescono a sopravvivere.

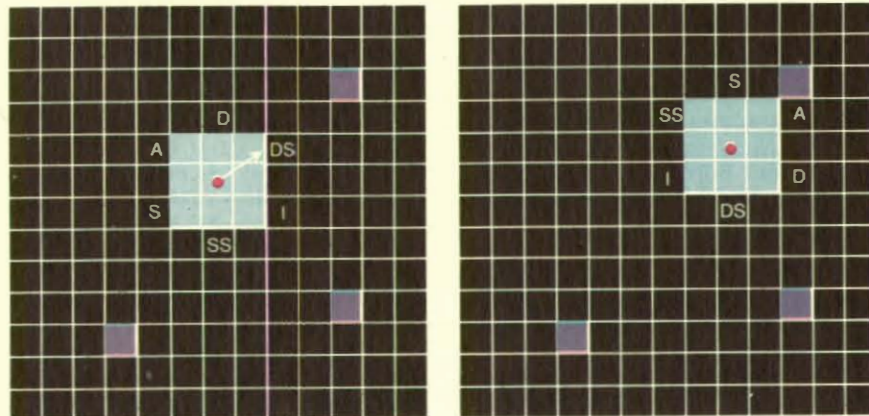
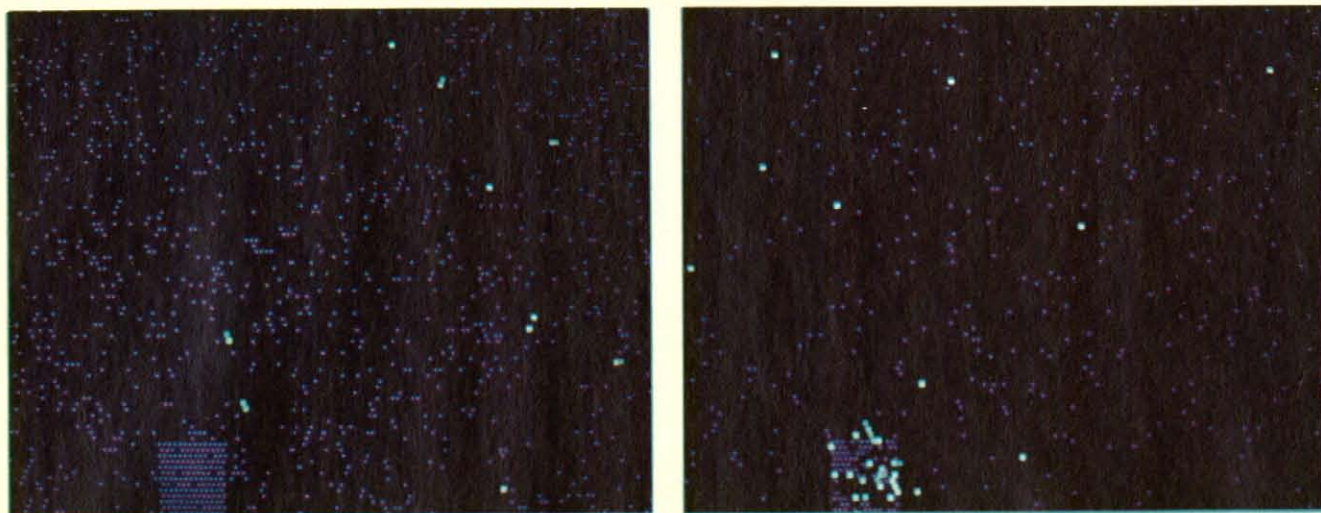
Le generazioni si succedono ogni minuto circa. Questa lotta in miniatura per la vita e la morte garantisce uno spettacolo avvincente, ma il dramma si fa molto più intenso dopo diversi minuti, quando l'osservatore si accorge che alcuni dei microbi hanno iniziato a comportarsi in modo diverso. Dopo qualche minuto, ci sono microbi che fanno acrobazie. Dopo 20 minuti o più alcuni microbi scivolano, almeno per brevi distanze. Questi microbi sembrano avere un successo molto maggiore dei loro antenati vaganti. In effetti, prolificano prima ancora che uno se ne renda conto.

A tempo debito, si sviluppano microbi «incrociatori» che per lo più si muovono in avanti, girando solo di tanto in tanto. Questo significa che si muovono quasi sempre verso popolazioni più dense di deliziosi batteri viola. Questo comportamento, una volta determinatosi in alcuni singoli microbi, arriva a dominare l'intera popolazione, dato che gli incrociatori finiscono per impadronirsi della maggior parte del cibo.

Anche se gli incrociatori costituiscono una specie evoluta, ci sono ancora delle variazioni all'interno della loro popolazione. Alcuni incrociatori, per esempio, girano più spesso a destra che a sinistra, mentre altri preferiscono le svolte a sinistra. Ogni tanto, naturalmente, ci sono anche passi indietro: qualche incrociatore genera discendenti incapaci di adattarsi. Una malattia genetica abbastanza diffusa è la «roteosi», in preda alla quale un microbo gira troppo spesso in una sola direzione. Le sfortunate creature che ne sono affette di solito muoiono prima di aver conosciuto la gioia della scissione.

È abbastanza interessante osservare l'emergere della specie degli incrociatori, ma il programma di Palmiter offre di più. Che cosa avviene nel caso di una variazione dell'ambiente? A questa domanda si può rispondere facendo girare Evoluzione simulata in una modalità nella quale lo schermo appare del tutto identico alla situazione precedente, tranne che per la presenza di una zona particolarmente ricca di batteri nell'angolo in basso a sinistra. In quella zona rigogliosa, che Palmiter chiama Giardino dell'Eden, il rifornimento di batteri avviene a una velocità molto più elevata del normale (*si veda l'illustrazione nella pagina seguente*).

Con il passare delle generazioni, gli incrociatori si evolvono come prima; ma all'interno del Giardino dell'Eden avviene qualcosa di molto diverso. I pochi microbi vaganti che si sono fortunosamente imbattuti nel banchetto di batteri sono prontamente ricompensati per la loro mancanza di un metodo di alimentazione organizzato. Possono ondeggiare quanto vogliono: il cibo continua a circondarli.

A mano a mano, però, che il cibo diventa più scarso nel Giardino dell'Eden, comincia a farsi sentire una lieve pressione ambientale. Vagare e ondeggiare non sono più strategie valide, ed è a questo punto che fanno la loro comparsa i «roteanti». Quello che solitamente costituisce un disastroso difetto genetico diventa invece un vantaggio in un Giardi-

Il Giardino dell'Eden (a sinistra) *favorisce l'evoluzione di «roteanti»* (a destra)

no dell'Eden sovrappopolato. In effetti, con il passare del tempo i microbi che manifestano una forte tendenza a girare in una sola direzione finiscono per prevalere nel giardino. La ragione è ovvia: un microbo che gira spesso nella stessa direzione, per esempio a destra, tende a rimanere nel Giardino dell'Eden significativamente più a lungo dei suoi antenati vaganti.

In qualche ora al massimo, il Giardino dell'Eden è popolato quasi unicamente da roteanti altamente specializzati, che si potrebbero anche chiamare orbitanti nervosi. Essi seguono una particolare orbita per molti cicli e poi si spostano improvvisamente di un quadratino e ripetono l'orbita, spazzando via batteri a ogni spostamento.

Il programma Evoluzione simulata è un modello valido dell'evoluzione biologica? Solo in un senso molto limitato. È vero che mostra come un ambiente possa favorire la comparsa di certe variazioni nella prole, variazioni che portano infine alla formazione di nuove specie, ma le somiglianze non vanno oltre. Una volta che sono venute alla luce una o due specie stabili di microbi, non accade più nient'altro. Che cosa occorrerebbe per poter realizzare il sogno di Dawkins di una continua evoluzione generata dal calcolatore? Forse niente di meno che un universo in miniatura nel calcolatore.

I lettori che vogliano studiare l'argomento possono ordinare una copia di Evoluzione simulata (al costo di 39,95 dollari: il nome inglese del programma è *Simulated Evolution*) a Life Science Associates, una piccola azienda che si occupa di software didattico. L'indirizzo è 1 Fenimore Road, Bayport, New York 11705. Il programma è stato scritto per i PC IBM e calcolatori compatibili, e viene fornito completo di un elaborato manuale. Per i programmatori relativamente avanzati che preferiscono scriversi la propria versione di Evoluzione simulata, fornirò ora una descrizione di BUGS, il nome che ho dato a una versione semplificata del programma.

Un microbo di BUGS può essere rappresentato sullo schermo da un quadratino di lato pari a tre pixel. Nell'illustrazione di pagina 94 si può vedere come si muove questo microbo. Una semplice tabella specifica il modo in cui le coordinate del pixel centrale di un microbo cambiano a seconda della direzione verso cui il microbo è orientato. La tabella contiene due vettori di sei elementi, *xmossa* e *ymossa*.

dir	0	1	2	3	4	5
xmossa	0	2	2	0	-2	-2
ymossa	2	1	-1	-2	-1	1

La direzione verso cui punta un microbo (rispetto allo schermo del calcolatore) è data dal valore della variabile *dir*. I numeri corrispondenti di *xmossa* e *ymossa* indicano di quanti pixel vada di conseguenza spostato il microbo in una sola mossa sullo schermo, rispettivamente nella direzione orizzontale e verticale. Se, per esempio, un microbo punta nella direzione 2, deve essere spostato a destra di due pixel e verso il basso di un pixel, in quanto $xmossa(2) = 2$ e $ymossa(2) = -1$.

BUGS determina la direzione del moto di ciascuna delle creature di sua competenza consultando una formula basata sul codice genetico di ciascun microbo, che è contenuto in una matrice a due dimensioni chiamata *gene*. L'elemento *gene*(k,j) contiene il valore del j-esimo gene del k-esimo microbo. Ciascun valore dei geni, nella formula, compare come esponente di 2, per non avere a che fare con numeri negativi. Per trovare la probabilità che un microbo si muova nella direzione d, si divide 2 elevato al valore del gene d per la somma di 2 elevato a ciascuno dei valori dei geni. Per esempio, la probabilità che il successivo movimento del microbo sia una curva stretta verso sinistra è data dalla divisione di 2^{SS} per la somma $2^A+2^D+2^{DS}+2^I+$ $+2^{SS}+2^S$.

In questo modo, BUGS calcola sei probabilità per ciascuno dei sei possibili movimenti. La somma delle probabilità è naturalmente uguale a 1. Si può pensare alle probabilità come a sei diversi domini che, insieme, abbracciano l'intero intervallo da 0 a 1. In altri termini, rappresentando con i simboli da p_0 a p_5 le probabilità delle sei differenti direzioni di moto, il dominio 0 è dato dall'intervallo da 0 a p_0, il dominio 1 dall'intervallo da p_0 a p_0+p_1, il dominio 2 dall'intervallo da p_0+p_1 a $p_0+p_1+p_2$ e così via.

A ciascun ciclo, BUGS determina poi un nuovo valore per la direzione del moto di un particolare microbo scegliendo a caso un numero tra 0 e 1, guardando in quale dominio esso venga a cadere e assegnando il numero del dominio alla variabile *gira*. In questo schema, allora, *gira* è uguale a 0 per A, a 1 per D, a 2 per DS, a 3 per I, a 4 per SS e a 5 per S.

Poche semplici istruzioni completano l'algoritmo del moto:

dir ← *dir* + *gira*(mod 6)
microbox(k) ← *microbox*(k) +
+ *xmossa*(*dir*)
microboy(k) ← *microboy*(k) +
+ *ymossa*(*dir*)

Le matrici *microbox*(k) e *microboy*(k) contengono le attuali coordinate del k-esimo microbo.

Nella prima riga, viene modificata la direzione attuale *dir* sommando il risultato della scelta casuale contenuta nella variabile *gira*. La somma deve essere modulare. Per esempio, se *dir*=5 (cioè il microbo si dirige in alto a sinistra) e *gira*=2 (cioè deve fare una curva stretta verso destra), il nuovo valore di *dir* sarà 5+2(mod 6)=1, e il movimento successivo del microbo è verso l'alto a destra.

BUGS deve muovere tutti i microbi secondo questa formula, controllando a ciascun passo se un microbo ha colpito una barriera o è finito su un batterio. Inoltre, deve registrare l'età e il rifornimento di energia di ciascun microbo, per stabilire se un determinato microbo debba estinguersi o possa riprodursi. Quando un microbo è pronto per la scissione, il programma sostituisce semplicemente il vecchio microbo con due nuovi nella stessa collocazione. Questi ereditano i valori genetici del vecchio microbo, tranne per il fatto che un valore scelto a caso è aumentato di una certa quantità in uno dei due discendenti e un altro valore scelto a caso è diminuito della stessa quantità nell'altro.

Il programma SLO GRO per la generazione di frattali, descritto nella rubrica dello scorso febbraio, ha suscitato grande entusiasmo. Il perdurare dell'interesse per i frattali, di qualunque forma essi siano, è testimoniato da una vera valanga di lettere. Scientificamente, il programma può essere descritto come un simulatore dell'aggregazione per diffusione limitata, un fenomeno che vediamo nella formazione di certi minerali, nella placcatura elettrolitica dei metalli e anche nell'accumularsi della fuliggine.

Le indicazioni date a proposito di SLO GRO sono risultate sufficientemente semplici per molti lettori, e molti in effetti le hanno seguite. L'algoritmo di base comporta l'inserimento in un cerchio, a partire da un punto a caso sulla circonferenza del cerchio stesso, di una «particella» che si muove in modo casuale. Quando la particella entra in contatto con un suo simile stazionario, interrompe anch'essa il suo movimento, producendo così un'aggregazione di particelle. Il programma era facile da scrivere ma alcuni l'hanno trovato piuttosto noioso da vedere al lavoro. Perché dovevano perdere tempo a osservare un punto di luce che sembrava vagare per un'eternità? Fatto sta che numerosi lettori hanno pensato di cambiare l'algoritmo per accelerarne l'azione.

Edward H. Kidera IV di Columbia (Maryland) ha ottenuto una discreta accelerazione partendo da un cerchio di piccole dimensioni e aumentando costantemente il suo raggio al crescere dell'aggregazione. Molti lettori hanno anche avanzato dei suggerimenti per accelerare la verifica dell'avvenuto contatto con un ammasso di particelle. Tale verifica prevede il confronto dei pixel adiacenti alla particella con le posizioni registrate di ogni particella dell'ammasso in crescita.

Quella che segue è la proposta fatta, a questo riguardo, da Ronald C. Read, dell'Università di Waterloo nell'Ontario. «Per coloro che usano il BASIC (come certamente fanno molti dei suoi lettori) c'è un modo molto più semplice, quello di usare il comando POINT del BASIC per stabilire se al pixel in questione è stato attribuito un colore. In effetti, in questo modo, si usa lo schermo come dispositivo di immagazzinamento.»

Il più impaziente di tutti era William H. Pratt di State College (Pennsylvania). Perché far vagare la particella in modo del tutto casuale? Perché non darle direttamente una posizione casuale in prossimità dell'ammasso? Pratt è rimasto però molto stupito nello scoprire che il suo ammasso non somigliava per niente alle illustrazioni dello scorso febbraio: era sì frastagliato ai margini, ma anche più solido: una creatura dall'aspetto del tutto diverso.

Senza saperlo, Pratt stava giocando con quello che è noto come modello di crescita di Richardson, uno strumento di ricerca molto amato da un gruppo di matematici definito scherzosamente «la mafia delle particelle». Questi ricercatori, alcuni dei quali lavorano all'Università del Wisconsin a Madison, hanno studiato per più di un decennio una grande varietà di modelli di crescita. Spero di poter dare, in uno dei prossimi articoli, un resoconto della visita che ho recentemente fatto a Madison.

Nella rubrica di marzo mi sono occupato dei rompicapo di persone: rompicapo logici che si possono risolvere solo pensando a ciò che altre persone stanno pensando. Un'intera classe di questi rompicapo era rappresentata da tre filosofi che si svegliavano da un pisolino fatto sotto un albero. Ciascun filosofo notava che sulla fronte degli altri due c'era quello che sembrava proprio un «regalino» di un uccello. Solo mentre stava già ridendo come gli altri due, il più saggio di loro si era reso conto che anche la sua fronte era sporca. Come era arrivato a questa deduzione?

Non avevo immaginato, a differenza di James D. Klein di College Place (Washington), che esiste quello che potremmo chiamare un rompicapo dei due filosofi. Klein ha sottoposto ai suoi figli la storia di due lavoratori che cadono a terra da un'impalcatura. Nella caduta, nessuno dei due si fa male, ma uno si sporca la faccia. Perché il lavoratore con la faccia pulita corre a lavarsi mentre quello con la faccia sporca torna semplicemente a lavorare? Scrive Klein: «È interessante sentirli pensare a voce alta e vedere i loro occhi quando viene a galla la soluzione».

Un altro rompicapo di persone era preso a prestito dal libro di Dennis Shasha *The Puzzling Adventures of Dr. Ecco*. In questo indovinello, due generali del diciannovesimo secolo, al comando di armate separate da un rilievo montuoso, decidono di coordinare l'attacco contro il nemico inviandosi messaggi con dei piccioni viaggiatori. Ma quale messaggio usare? Se il primo generale invia il messaggio «Attacchiamo all'alba», deve attendere un messaggio di conferma da parte del secondo generale. E se uno dei piccioni non arriva mai dall'altra parte? E se anche tutti e due i piccioni arrivano a destinazione, come fa il secondo generale a sapere che la sua conferma è stata ricevuta? Si direbbe che sia inevitabile cadere in una serie infinita di messaggi.

Il dilemma dei generali ha fatto venire in mente a Warner Clements di Beverly Hills (California) un lavoro teatrale poco noto in cui un doppio agente di mezza tacca fa la spola tra due nazioni nemiche. La storia inizia quando l'agente viene a sapere che il paese A ha violato il codice militare segreto del paese B. L'agente va in B per vendere l'informazione ai servizi segreti di quel paese. «Lo sappiamo già», gli rispondono. L'agente rimane dapprima scoraggiato ma poi si rende conto che può rivelare ai servizi segreti di A che anche B ha violato il loro codice. Questi a loro volta gli rispondono: «Sappiamo che i B hanno violato il nostro codice. Gli stiamo inviando informazioni false!» L'agente si precipita di nuovo nel paese B: «Vi rendete conto che gli A sanno che avete violato il loro codice?» «Oh, sì», gli rispondono. L'agente torna dagli A per metterli al corrente della situazione, e così via. Per quanto tempo, in teoria, dovrebbero continuare i viaggi avanti e indietro dell'agente, con un messaggio sempre più lungo a proposito di quello che sanno gli avversari? Il fatto che in questo rompicapo le fazioni militari siano in competizione invece che in collaborazione non rende più semplice la soluzione, anche perché la soluzione non esiste proprio.

Rompicapo di persone più realistici hanno a che fare con persone concrete in situazioni quotidiane, come quelle studiate dal defunto sociologo Erving Goffman. Ne avevo chiesto qualche esempio e ne ho ricevuti parecchi, tra cui uno da P. M. Cambeen, di Muiden, nei Paesi Bassi. Durante la seconda guerra mondiale, un ufficiale delle forze tedesche che occupavano il paese espresse a un abitante del luogo le sue perplessità sugli atteggiamenti e le opinioni degli olandesi. Gli fu risposto: «Gli olandesi hanno tre virtù: sono intelligenti, leali e favorevoli ai nazisti. Un olandese qualsiasi, però, ha solo due di queste virtù e l'opposto della terza». Mentre l'ufficiale ragionava sulle implicazioni logiche di questa affermazione, l'arguto olandese ebbe tutto il tempo di andarsene.

BIBLIOGRAFIA

DAWKINS RICHARD, *L'orologiaio cieco*, Rizzoli, Milano, 1988.

LANGTON CHRISTOPHER G. (a cura), *Artificial Life: The Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems Held September, 1987, in Los Alamos, New Mexico*. Addison-Wesley Publishing Company, Inc., 1989.